



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### A Higher-Order Approach to Ontology Evolution in Physics

**Citation for published version:**

Lehmann, J, Chan, M & Bundy, A 2013, 'A Higher-Order Approach to Ontology Evolution in Physics' Journal on Data Semantics, vol 2, no. 4, pp. 163-187. DOI: 10.1007/s13740-012-0016-7

**Digital Object Identifier (DOI):**

[10.1007/s13740-012-0016-7](https://doi.org/10.1007/s13740-012-0016-7)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Early version, also known as pre-print

**Published In:**

Journal on Data Semantics

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# A Higher Order Approach to Ontology Evolution in Physics

Jos Lehmann ·  
Michael Chan ·  
Alan Bundy

Received: date / Accepted: date

**Abstract** A Higher-Order Logical approach to ontology evolution is applied to examples in Physics. Based on this approach, a framework is proposed which includes a methodology for the formalization of ontology evolution in Higher-Order Logic and an implementation of this methodology in the theorem prover Isabelle. The proposed basic mechanisms for evolution are called *Ontology Repair Plans*. These operate on ontologies formalised as contexts, i.e., as multiple logical theories. In such a setting, ontologies may contradict one another or introduce redundancies with respect to one another, without any of them containing logical contradictions or redundancies. When, though, an inconsistency or a redundancy between two or more ontologies becomes explicit, it may be resolved by the application of an Ontology Repair Plan, as each plan compiles together a pattern for diagnosis and transformation rules for effecting a repair. The repair can combine the retraction and/or addition of axioms as well as the deeper modification of the language in which the ontology is represented.

**Keywords** Higher Order Logic, Ontology Evolution, Physics.

---

J. Lehmann  
University of Edinburgh, School of Informatics, 10 Crichton Street,  
Edinburgh EH8 9AB, United Kingdom (author's affiliation throughout  
GALILEO project)  
Mob: +44-79-48476419  
E-mail: jos.lehmann@gmail.com

M. Chan  
University of Edinburgh, School of Informatics, 10 Crichton Street,  
Edinburgh EH8 9AB, United Kingdom  
E-mail: mchan@inf.ed.ac.uk

A. Bundy  
University of Edinburgh, School of Informatics, 10 Crichton Street,  
Edinburgh EH8 9AB, United Kingdom  
E-mail: a.bundy@ed.ac.uk

## 1 Introduction

Artificial Intelligence (AI) and, more generally, Computer Science (CS) are presently faced with the challenge of how autonomous software can manipulate its own knowledge. Such knowledge is typically represented in an ontology that conceptualises the entities of the software's application domain and allows the software to reason about such entities at a higher level of abstraction than simply the level of data or information. Just like any abstract model, ontologies are limited representations of the world, which is dynamic and inherently complex. If autonomous systems are to feature any kind of robustness with respect to the dynamics and complexity of changing environments and goals, of communication acts and of new information, they must be able to autonomously update their own ontologies.

The literature on the subject of updating an ontology in the face of new information often uses the phrase *ontology evolution* and usually concentrates on how Description Logic (DL) axiomatic theories for Semantic Web applications or First Order Logic (FOL) theories need to retract axioms or modify entailments in order to maintain their coherence and consistency. Section 2 discusses how this type of work on ontology evolution is related to ours.

Based on this discussion, we present our approach to ontology evolution in Higher-Order Logic (HOL). This approach underlies a framework called GALILEO, for Guided Analysis of Logical Inconsistencies Leads to Evolved Ontologies. There are a few previous publications on GALILEO, e.g. (Bundy and Chan, 2008; Chan and Bundy, 2008; Chan et al., 2010). Here we go beyond those papers by firmly positioning our proposal with respect to the wider literature on ontology evolution and by presenting a stable integration, standardization, implementation and evaluation of the framework<sup>1</sup>.

GALILEO includes a methodology for the formalization of ontology evolution in higher-order Logic (HOL), as well as an implementation of this methodology in the automated theorem prover Isabelle. As discussed in more detail below, in GALILEO the term "ontology" is interpreted in an inclusive way, to encompass any representation of knowledge, usually as a logical theory. We have chosen a logic that is rich enough to represent, in a natural way, the object-level concepts and relationships of Physics as well as the meta-level concepts and relationships of ontology evolution processes. This argues for a logic such as HOL, which is polymorphic<sup>2</sup>, typed, and higher-order.

---

<sup>1</sup> Further workshop reports and the implementation of GALILEO can be downloaded from [http://dream.inf.ed.ac.uk/projects/ontology\\_evolution/](http://dream.inf.ed.ac.uk/projects/ontology_evolution/)

<sup>2</sup> *Polymorphic* means that some terms of the language can have more than one type.

In GALILEO, the basic evolution-mechanisms are called *Ontology Repair Plans* (ORPs). Each ORP compiles together a pattern for diagnosis of conflicts between ontologies and transformation rules for effecting a repair. For both development and testing, we rely on examples from Physics, as advances in this field may naturally be modelled as cases of ontology evolution and they are usually well documented. Physicists revise predictive theories when confronted with conflicting experimental evidence. Therefore, the ORPs typically assume there is one ontology representing a predictive theory and other ontologies representing an experimental or observational set-up for that theory. When the experimental ontology contains a theorem that contradicts or is redundant with respect to a theorem of the theoretical ontology, an ORP is triggered and amends the two ontologies so that the conflict is resolved. The development methodology of GALILEO revolves around the selection of initial ontologies by the collection, analysis, formalisation, implementation and testing of appropriate case studies in the history of Physics and of the ORPs inspired by them.

The initial ontologies are formalized and implemented as contexts, i.e., as multiple logical theories, which are isolated from other theories and which may even use their own local language to describe the entities in the domain. In such a setting, if the ontologies share the same language, they may *explicitly* contradict one another or introduce redundancies with respect to one another, without any of them containing logical contradictions or redundancies. On the other hand, if the ontologies do not share the same language they may *implicitly* contradict one another or introduce redundancies with respect to one another without producing an *explicit* logical contradiction or a redundancy at the global level. When, though, two or more such ontologies are *bridged*, i.e., mapped or aligned through a third one which aligns them and resolves their differences, they will be able to share axioms. This will allow the proof of contradictory facts or make some facts redundant, thus affecting the robustness of the system. In both these types of situation (i.e., shared or non shared language) we say that the ontologies are *locally consistent or parsimonious* but *globally inconsistent or redundant*. When the global inconsistency or redundancy becomes explicit the application of an ORP may resolve the conflict between the two ontologies.

Shortly, we investigate in Higher-Order Logic the problem of ontology evolution from the perspective of automating the mechanisms to repair locally consistent or parsimonious but globally inconsistent or redundant ontologies.

The empirical part of our methodology supports the long-term objectives of the definition of a theory of ontology evolution, as well as of an evolution calculus. In this article we test the following hypothesis:

*Automated ontology evolution via ORPs is computationally feasible and can account for the kinds of on-*

*tology evolution that are observed in human problem solving in the Physics domain. We show that desirable properties, e.g., coverage, efficiency, maintainability, high quality of the repairs, can be achieved.*

The empirical results provided by the evolutions proposed by GALILEO allow for the evaluation of all these desirable properties. In particular, the results allow us to conclude that the generality of a Higher-Order Logical approach to ontology evolution is beneficial for the quality of the repairs, in terms of the meaningfulness of the proposed evolutions.

Section 2 discusses related work and makes the case for a higher order approach. Section 3 describes the adopted methodology. Sections 4, 5, 6 present three ORPs and applications for each them. These three sections are quite dense in formal content, they describe in detail the formal apparatus of our proposal as well as a number of models of case studies to which our ORPs apply. Section 7 illustrates the implementation of GALILEO. Section 8 discusses the evaluation. Section 9 discusses further work. Section 10 draws some conclusions.

## 2 Discussion of related work

The literature on the subject of updating an ontology in the face of new information often uses the phrase *ontology evolution* and usually concentrates on how DL ontologies for Semantic Web applications or FOL theories need to evolve, either to maintain their own coherence and consistency (ontology debugging, in a broad sense) or to establish a relationship with other ontologies (ontology alignment, in a broad sense).

Debugging yields notions like incoherence and inconsistency diagnosis and repair (Haase et al., 2005; Kalyanpur et al., 2006b; Ji et al., 2009; Lam et al., 2008), belief revision (Flouris, 2006), conservative extensions (Ghilardi et al., 2006). There are also pragmatic approaches such as (Lösch et al., 2009) that, in analogy with database technology, tackle the problem of updating an ontology by means of domain-specific knowledge about temporal changes. Such pragmatic solutions fall outside the scope of this discussion, which concentrates on domain-independent and logic-based methods for ontology evolution.

Ontology alignment, on the other hand, yields notions like matching (Doan et al., 2004; Giunchiglia and Shvaiko, 2004), mapping, bridging, contextualisation (Bouquet et al., 2004; Stuckenschmidt et al., 2004), distribution (Borgida and Serafini, 2003; Stuckenschmidt, 2006; Homola, 2007).

Both debugging and alignment impact ontology evolution, but they are somewhat orthogonal to each other. Typically, approaches to alignment examine whether some situations yielded by aligning or integrating ontologies is problematic for some desirable properties of DL, and often a

$$\begin{aligned}
\text{sig}(\mathcal{T}_{uni}) &::= \{ \\
& \text{Person}, \text{Student}, \text{Employee}, \dots \} & (1) \\
\mathcal{T}_{uni} &::= \{ \\
& \text{Student} \sqsubseteq \text{Person}, & (2) \\
& \text{Employee} \sqsubseteq \text{Person}, & (3) \\
& \text{Employee} \sqsubseteq \neg \text{Student}, & (4) \\
& \text{PhDStudent} \sqsubseteq \text{Student}, & (5) \\
& \text{PhDStudent} \sqsubseteq \text{Employee} \} & (6) \\
\mathcal{A}_{uni} &::= \{ \\
& \text{bruce} : \text{PhDStudent} \} & (7)
\end{aligned}$$

Fig. 1: *Incoherent and inconsistent DL ontology of members of a university.* Note the following DL conventions:  $A \sqsubseteq B$  is the DL notation for the first-order logic formula  $\forall x.A(x) \rightarrow B(x)$ ;  $i : A$  is the DL notation for  $i$  is an instance of  $A$ ; a DL ontology consists of a TBox,  $\mathcal{T}$ , in which concepts are specified using the signature elements of  $\text{sig}(\mathcal{T})$ , and of an ABox,  $\mathcal{A}$  where assertions are made about the individuals using the terminology; the axiomatic theory of members of a university formed by the ordered couple  $(\mathcal{T}_{uni}, \mathcal{A}_{uni})$  is a knowledge base. In DL terms (Flouris et al., 2006),  $(\mathcal{T}_{uni}, \mathcal{A}_{uni})$  is *inconsistent*, i.e., it has no models, because axioms (7, 5) allow to conclude that *bruce* is a *Student* while axioms (7, 6, 4) allow to conclude that *bruce* is not a *Student*. The source of the inconsistency lies in  $\mathcal{T}_{uni}$ , which is *incoherent*, i.e., it contains the *unsatisfiable* concept *PhDStudent*. A concept is unsatisfiable if it is mapped to the empty set in all models of the ontology.

modular approach is proposed to avoid those situations. As opposed to this, the literature on ontology debugging, typically tries to pin down how to diagnose situations that are already acknowledged as problematic (e.g. inconsistencies, redundancies, etc.), or how to repair ontologies that suffer such problems.

Although our proposal in this article contains elements of alignment (as mentioned, ontologies are formalized as HOL contexts), our main focus is on diagnosis and repair. Therefore we do not delve into the literature on alignment, but discuss relevant examples of the literature on debugging.

A standard example used in the DL literature about incoherence and inconsistency diagnosis and repair is shown in Figure 1. In DL terms (Flouris et al., 2006), the axiomatic theory of members of a university formed by the ordered couple  $(\mathcal{T}_{uni}, \mathcal{A}_{uni})$  is *inconsistent*, i.e., it has no models, because it allows to conclude a contradictory fact, i.e., that *bruce* is both a *Student* and not a *Student*. The inconsistency is due to the *incoherence* of  $\mathcal{T}_{uni}$ , in particular to the concept *PhDStudent* which is *unsatisfiable*, because it is mapped to the empty set in all models of the ontology.

In order to diagnose the unsatisfiable concept and repair the incoherence and the inconsistency, DL approaches such as (Haase et al., 2005) try to identify the axiom that, if removed, would allow to re-establish coherence and consistency. This is done by algorithms that compute so-called *Minimal Unsatisfiability-Preserving Subsets* of the ontology (MUPS). For instance, starting from a subset of the ontology containing one axiom, the algorithm adds to the subset, one at the time, the axioms that are *connected* to it, i.e., that share type symbols with it (e.g., axioms (4, 5)), until the subset is unsatisfiable. In Figure 1 the only MUPS of  $(\mathcal{T}_{uni}, \mathcal{A}_{uni})$  contains axioms (4, 5, 6, 7). A judgement call is therefore required to decide which axiom to remove. This decision procedure may increase in complexity in scenarios in which it is more than one group of axioms that independently make a given concept unsatisfiable, thus multiplying the number of MUPS.

Even in cases where only one MUPS is returned, the axiomatic structure of the ontology may be entangled and make axiom-removal simply infeasible. Consider for instance the version of  $(\mathcal{T}_{uni}, \mathcal{A}_{uni})$  given below (where  $/$  is set subtraction) in which axioms (4, 5) are substituted by one single axiom (8):

$$\begin{aligned}
\mathcal{T}_{uni}^* &::= \mathcal{T}_{uni} / \{ \text{Employee} \sqsubseteq \text{Person}, \\
& \text{Employee} \sqsubseteq \neg \text{Student} \} \\
& \cup \{ \text{Employee} \sqsubseteq \text{Person} \sqcap \neg \text{Student} \} & (8)
\end{aligned}$$

Removing axiom (8) would affect the entailment structure of  $\mathcal{T}_{uni}^*$  beyond what is needed and make the repair harmful, because an instance of *Employee* would no more qualify as an instance of *Person*. This has motivated approaches such as (Lam et al., 2008; Kalyanpur et al., 2006a; Schlobach and Cornet, 2003) that try to refine both the diagnostic and the repair operations, in order to allow for more surgical removals or even axiom rewriting. In such approaches the diagnosis usually applies minimal unsatisfiability to abstract versions of the axioms at hand. For instance, axiom (8) would have to be broken up along its conjunction, i.e., reduced to (4, 5), so that a separate unsatisfiability test can be run on each conjunct. A problematic aspect of this approach to diagnosis is that different types of formulae and connectives need different types of abstraction rules and, of course, the complexity of the rules increases with the complexity of the axioms. Abstraction may also become a problem when repairing, because the user needs to be able to relate back the abstracted axioms to the original ones in order to proceed to their removal, and this is not always a straightforward task.

An additional type of incoherence and inconsistency diagnosis and repair is presented in (Ji et al., 2009). This applies a MUPS-based, axiom-removal approach to the case of multiple aligned ontologies, so-called ontology networks, in particular to the diagnosis and repair of networked ontologies that are locally consistent and globally inconsis-

tent. Figure 2 shows an example of a coherent and consistent version of  $(\mathcal{T}_{uni}, \mathcal{A}_{uni})$ ,  $(\mathcal{T}_{uni}^{**}, \mathcal{A}_{uni}^{**})$ , that is mapped onto another coherent and consistent ontology  $(\mathcal{T}_{com}, \mathcal{A}_{com})$  for company members. The network formed by these two ontologies through the mapping  $Map(\mathcal{T}_{uni}^{**}, \mathcal{T}_{com})$  is incoherent and inconsistent, as it allows to prove again that *bruce* both is and is not a *Student*. Coherence and consistency are re-established by repairing the mapping.

A non-DL-based approach to problems of ontology evolution is implemented in the Ontology Repair System (ORS) (McNeill and Bundy, 2007). ORS addresses the problem of ontology alignment in a multi-agent planning environment: a problem which it is essential to solve in order to realise the vision of the Semantic Web. In the ORS environment some agents offer services and others require these services. Each agent represents these services with planning action rules similar to those of the Stanford Research Institute Problem Solver (STRIPS), i.e., with preconditions and effects written in a first-order logical language, namely, in the Knowledge Interchange Format (KIF). It is inevitable with any sufficiently large agent community that there will be differences between their ontologies, even when there have been attempts to standardize them. The purpose of ORS is to identify and repair these ontological mismatches at run time, since compile-time alignment is unrealistic in this scenario. Moreover, this has to be done without full access to the other agents' ontologies, since full access is also unrealistic. ORS's ontology repair differs from the more common ontology matching in being aimed at identifying and correcting errors in a single ontology rather than constructing a mapping between two ontologies. Ontology repair is also done automatically, dynamically and without access to the ontologies of other agents, whereas typically ontology matching has some manual element, is done statically and with full access to both ontologies. The ORS ontology repair operations consist mostly of syntactic manipulations of the underlying logical representation, e.g., the number or order of the arguments of a function may be changed; a function may be divided into two or more, or two or more may be merged into one. The only belief revision operations available to ORS are to add or remove a precondition of an action rule. ORS has been successfully evaluated by showing that it emulates a high percentage of their owners' manual modifications to third party ontologies taken from the KIF and planning communities.

In general, we think that the approaches described above, particularly the DL-based ones, have proven insufficient to address the study of the *automation* of ontology evolution, because of a number of circumstances:

1. Most proposed approaches depend on users' instructions, which does not include ontology evolution performed *at runtime*, for instance, by agents in heterogeneous environments.

$$\mathcal{T}_{uni}^{**} ::= \mathcal{T}_{uni} / \{Employee \sqsubseteq \neg Student\}$$

$$\mathcal{A}_{uni}^{**} ::= \mathcal{A}_{uni}$$

$$sig(\mathcal{T}_{com}) ::= \{ Person, Trainee, Staff, \dots \} \quad (9)$$

$$\mathcal{T}_{com} ::= \{ Trainee \sqsubseteq Person, \quad (10)$$

$$Staff \sqsubseteq Person, \quad (11)$$

$$Staff \sqsubseteq \neg Trainee \} \quad (12)$$

$$Map(\mathcal{T}_{uni}^{**}, \mathcal{T}_{com}) ::= \{ \langle 1, Student, Trainee \rangle, \quad (13)$$

$$\langle 2, Employee, Staff \rangle \} \quad (14)$$

Fig. 2: *Locally coherent and consistent and globally incoherent and inconsistent network of DL ontologies of members of a university and of a company.* Note the following convention:  $\langle id, t_i, t_j \rangle$  means that term  $t_i$  is mapped onto term  $t_j$ .  $(\mathcal{T}_{uni}^{**}, \mathcal{A}_{uni}^{**})$  is a coherent and consistent version of  $(\mathcal{T}_{uni}, \mathcal{A}_{uni})$  and is mapped onto another coherent and consistent ontology  $(\mathcal{T}_{com}, \mathcal{A}_{com})$  for company members. The network formed by  $(\mathcal{T}_{uni}^{**}, \mathcal{A}_{uni}^{**})$  and  $(\mathcal{T}_{com}, \mathcal{A}_{com})$  through the mapping  $Map(\mathcal{T}_{uni}^{**}, \mathcal{T}_{com})$  is incoherent and inconsistent, as it allows to prove again that *bruce* both is and is not a *Student*. Coherence and consistency are re-established by removing lines (13) or (14) from the mapping.

2. Automated ontology repair systems, e.g., the Semantic Web Ontology editor (SWOOP) (Kalyanpur et al., 2006b), or even those that take heterogeneity into account like the Repair and Diagnose Ontology Network (RADON) (Ji et al., 2009), focus mostly on retracting axioms and modifying entailments. Systems like (Lam et al., 2008) that go beyond axiom retraction and that try to enable refined repair operations seem to be doing so on a limited logical basis and they still have to rely on users' judgements in order to evaluate the helpfulness and harmfulness of repairs. Furthermore, such systems do not support deeper syntactic manipulations, e.g., changes of the signature of the language.
3. The focus on DL or FOL ontologies does not allow for a sufficiently general analysis and resolution of ontological faults. The limited expressivity of FOL, let alone fragments of it such as DL, constitutes a limit on the modelling of ontology evolution. Without the means to quantify over and to reason about the predicates, it is virtually impossible to formalise and automate *sufficiently generic* ontology evolution patterns.

In order to deal with the first two limitations found in the literature, we investigate the problem of ontology evolution from the perspective of maximizing the potential for

automation of syntactic operations more complex than axiom retraction only. In particular, *we focus on repair operations that can evolve the language of the ontology*. To this end we concentrate on the repair mechanisms for *locally consistent or parsimonious but globally inconsistent or redundant* ontologies. As mentioned in Section 1, we assume that information about the same entities is distributed in a system across different ontologies or contexts. If the contexts share the same language, they may *explicitly* contradict one another or introduce redundancies with respect to one another, without any of them containing logical contradictions or redundancies. On the other hand, if the contexts do not share the same language they may *implicitly* contradict one another or introduce redundancies with respect to one another without producing an *explicit* logical contradiction or a redundancy at the global level. When two such contexts are bridged, i.e., mapped through a third one which resolves their differences, they will be able to communicate and prove facts in terms of each other. In both these types of situation (i.e., shared or non shared language) a multiple-ontologies approach has inherent advantages when dealing with conflicts. In particular in our ORP-based approach, multiple ontologies are required in order to look at the structure of proofs of inconsistencies, i.e., at how conflicts arise between particular sets of theorems that match our diagnostic patterns (called *triggers* or *trigger formulae* in GALILEO). In a consistent ontology it is significant that a particular trigger formula is provable. In an inconsistent one, all trigger formulae would be provable, so we would learn nothing. Since the collection of trigger formulae is itself inconsistent, then we need to divide them into consistent subsets and allocate them between consistent ontologies in order for the trigger process to be meaningful and for our ORP technique to work.

As mentioned above we focus on an ORPs method rather than on the more widely used MUPS method because we are interested in signature changes, not (just) axiom deletion. MUPS-based approaches are well suited to axiom deletion as they identify exactly the axioms that are deletion candidates. But the ORP triggers are better aimed at signature change because they identify precisely the concepts whose representation needs change. This is because ORP's diagnoses identify the point at which the implicit inconsistency becomes explicit, whereas MUPS identify an underlying cause, if this cause is faulty axioms rather than faulty concepts. ORPs open up the new possibility of many kinds of syntactical manipulations (e.g., splitting a function, changing its arity, etc.) to re-establish global consistency. Note incidentally that a MUPS-based approach like (Ji et al., 2009), as well as most belief-revision-based approaches, cannot fully exploit the potential inherent in the multiple-ontologies case. They too are better suited to the retraction of axioms and mappings, or they attempt the intricate modification of the logical structure of an ontology on a limited logical basis.

Also, such approaches do not consider the modification of the ontology's signature. Certainly, ontology evolution requires changes to both the axioms and the signature. Our ORPs focus more on signature changes than axiom changes because our empirical studies identified these as more fundamental. The ORPs do, though, also make axiom changes. If intricate changes were proved necessary, they could be encompassed within our richer HOL representation. In that sense, there is a degree of complementarity between our work and traditional DL approaches.

For what concerns again the second, as well as the third limitation found in the literature, i.e., the inherent lack of generality, the GALILEO system bases automated ontology evolution on the use of HOL. In order to do more than retracting axioms and to achieve the subtler diagnosis and repairs mentioned above, a meta-logic more expressive than DL or FOL is required. Modelling ORPs in HOL allows for the existential quantification over ontologies and functions, which is useful in domains containing many ontologies/theories, e.g., natural sciences and general real-world semantics. Moreover, the polymorphism of symbols employed in ORPs permits their high generality. HOL formulae can abstract over types, number of arguments, etc. therefore using HOL as a meta-logic for ontology evolution in the long run facilitates the formulation of *declarative theories* of ontology evolution – as opposed to the *algorithms* usually found in the DL literature. In GALILEO, meta-level inference consists only of matching object-level formulae against the trigger patterns, and then applying the repairs dictated by the repair patterns. These inferences are entirely automatic. By using HOL patterns we can describe ontology evolution in a succinct and transparent way that might be disguised in an algorithmic presentation – especially if the lack of expressivity in the algorithm meant that it had to have separate cases to deal with different types, arities and orders. Note also that the repaired ontologies should not usually be a logical consequence of the originals, since the originals are often globally inconsistent and the repaired ones should not be. The next section provides more detail on using HOL to represent ontologies and their evolution.

An additional consideration for preferring HOL over DL or FOL is its expressivity as an object language, i.e., the possibility of representing directly, rather than as roles in concepts, mathematical knowledge. Many concepts in Physics are naturally higher-order. For instance, differential and integral calculus are 2nd-order functions; the distribution of matter in a galaxy is represented as a function from space-time to mass; the orbit of a planet is represented as a function from space-time to velocity; etc. It is only during reasoning within an object-level HOL theory that interaction is required with the theorem prover. It is unlikely that this problem could be cured by a change to a less expressive logic. We know of no decidable logic capable of carrying out the nec-

essary mathematical manipulations. Yet, having Mathematics as part of the representation allows to reason about the constraints underlying a given DL typology. Consider, for instance, a DL typology such as the Ontology of Astronomical Object Types (Cambresy et al., 2010). Here the roles that, for instance, represent measurements of brightness, polarization, period, velocity, mass etc. are each represented separately as properties of an object, which is classified based on the values (of combinations of) such properties found in a database. It is not possible, though, to represent the mathematical relations between these properties. This makes the representation less flexible and limits the possibility of reasoning about the typologies and adapt them to, for instance, progress in Physics, because it is not possible to define the properties underlying a typology as functions of one another.

In summary, the use of HOL in describing ORPs enables each ORP to diagnose and repair a wide range of faulty Physics theories, which may involve objects of different type and order. In less expressive logics this generality is not available, leading to the need for *ad hoc* mechanisms for different classes of example. It could be argued that, compared to DL approaches, automation is lost in the object-level reasoning within each physical theory, because, as we shall see from Section 7, the richness of reasoning in HOL requires human interaction. Nevertheless, this object-level reasoning is largely algebraic manipulation, of a kind that could not be represented in DL logics. So, a superficial contrast between decidable, totally automatic DL reasoning and this HOL reasoning is misleading. Section 8.4 discusses and illustrates the degree of automation we have achieved in GALILEO.

### 3 Evolution by Ontology Repair Plans and Contexts

At the heart of our approach to ontology evolution lies the notion of Ontology Repair Plan. These are generic combinations of diagnosis patterns and repair operations formulated in HOL that guide the evolution of an ontology. The diagnostic component takes as input two or more ontologies, and checks whether specific logical conditions (e.g., a contradiction or a redundancy) hold between the ontologies. In the positive, the repair component is triggered. By grouping these meta-level operations, ORPs trade-off completeness against reduction in search. The following subsections explain and define: the use of HOL for representing ontologies (Types, Terms, Type Inheritance, Free Variables); the notion of ontologies as contexts (Higher-Order Ontologies, Combining Ontologies, Local and Global Inconsistency, Ontology Fault); the use of multiple ontologies (Bridging Ontologies, Factorised Representation); the basic mechanisms for ontology evolution (Ontology Repair Plans).

#### 3.1 Representing ontologies and their evolution in HOL

*Ontology* originated in Philosophy, where it is the philosophical study of the nature of being, existence or reality. The word “ontology” has been adopted and adapted within Computing as meaning a formal representation of the concepts within a domain and the relationships between those concepts. Many ontologies consist only of a classification of the types of object or relationships within a domain, perhaps arranged hierarchically. Some consist of sets of RDF triples representing a relationship between two objects. Richer ontologies can consist of sets of formulae in a Description Logic, such as the Web Ontology Language OWL, or a First Order Logic, such as KIF. We can, thus, see that there are a range of interpretations of “ontology”:

- from a single description of the nature of reality to multiple, perhaps conflicting, descriptions of many smaller domains;
- from foundational (or upper) conceptualizations covering very general abstract notions, such as space, time, objects, processes, qualities, regions, etc., to domain-specific ones describing specific versions of such general notions.
- from a simple set of type declarations to a representation of world knowledge (where this knowledge is similar yet relatively more prone to changes than the knowledge encoded in foundational or domain ontologies);
- from a simple classification to a rich logic.

In GALILEO we have interpreted “ontology” in an inclusive way to encompass any representation of knowledge, usually as a logical theory. We have chosen a logic that is rich enough to represent, in a natural way, both the object-level concepts and relationships of Physics as well as the meta-level concepts and relationships of ontology evolution processes. As we will see, this argues for a polymorphic, typed, higher order logic for both purposes. *Polymorphic* means that some terms of the language can have more than one type. This will be achieved by including variables ranging over types in their type declarations.

The choice required a trade-off between inferential complexity and representational richness. In order not to burden ourselves with knowledge encoding problems, we have favoured representational richness. The price we have paid is that inference must often be interactive. When ontology evolution is much better understood, it will be time to increase the degree of automation by exploring the potential for representation of Physics and ontology evolution in a more restricted logic.

The formal definition<sup>3</sup> of our polymorphic, typed, higher order logic are given below.

<sup>3</sup> Definitions are denoted by the symbol ::= throughout.

**Definition 1 (Types)** *The following BNF describes polymorphic types:*

$$\mathcal{T} ::= \mathcal{V} \mid C \mid \mathcal{T} \Rightarrow \mathcal{T} \mid \mathcal{T} \times \mathcal{T} \mid \text{Set}(\mathcal{T})$$

where  $\mathcal{V}$  is the set of type variables,  $C$  the set of type constants;  $\tau \Rightarrow \tau'$  is the type of functions from objects of type  $\tau$  to objects of type  $\tau'$ ;  $\tau \times \tau'$  is the type of products from objects of type  $\tau$  to objects of type  $\tau'$ ; and,  $\text{Set}(\tau)$  is the type of sets of objects of type  $\tau$ . We will use the Greek letter  $\tau$ , possibly subscripted or primed, to range over types.

Note that the set of type constants,  $C$ , depends on the signature of the ontology.

Particular type constants that we use below are, for instance,  $\text{bool}$ , for the booleans, and  $\mathbb{R}$  for the real numbers.

**Definition 2 (Terms)** *The following BNF describes terms of the logic.*

$$T ::= V \mid C \mid T(T) \mid \lambda V.T$$

where  $V$  is the set of term variables,  $C$  the set of term constants. We will use the Roman letter  $t$ , possibly subscripted or primed, to range over terms. We will write  $t:\tau$  to declare term  $t$  to have type  $\tau$ .

Note that the set of constants,  $C$ , depends on the signature of the ontology.

We will assume that  $C$  contains: the truth values  $\top:\text{bool}$  (true) and  $\perp:\text{bool}$  (false); the standard logic connectives, such as  $\wedge:\text{bool} \Rightarrow \text{bool} \Rightarrow \text{bool}$ ; and the quantifiers, such as  $\exists:(\tau \Rightarrow \text{bool}) \Rightarrow \text{bool}$ . Most ontologies will also contain, for each type  $\tau$ , an equivalence relation  $=_\tau:\tau \Rightarrow \tau \Rightarrow \text{bool}$ , a partial order relation  $<_\tau:\tau \Rightarrow \tau \Rightarrow \text{bool}$  and additive and subtractive functions  $+_\tau, -_\tau:\tau \Rightarrow \tau \Rightarrow \tau$ . Note that several of these term constants are *polymorphic*, i.e., their types contain type variables. This allows us to overload these constants so that they apply to objects of many different types. This is essential in allowing us to formalise generic ontology repair plans that are applicable across a wide range of different areas of Physics.

**Definition 3 (Type Inheritance)** *Terms inherit their types according to the following two rules:*

$$\frac{t':\tau \Rightarrow \tau' \quad t:\tau}{t'(t):\tau'} \qquad \frac{x:\tau' \quad t:\tau}{\lambda x.t:\tau' \Rightarrow \tau}$$

**Definition 4 (Free Variables)** *The set,  $FV(t)$ , of free variables of a term,  $t$ , are defined recursively as follows:*

$$\begin{aligned} \forall x \in V. FV(x) &::= \{x\} \\ \forall t, t' \in T. FV(t'(t)) &::= FV(t') \cup FV(t) \\ \forall c \in C. FV(c) &::= \{\} \\ \forall t \in T, x \in V. FV(\lambda x.t) &::= FV(t) \setminus \{x\}. \end{aligned}$$

The above definitions of types and terms are in, so called, *curried form*, where all functions are regarded as unary. Curried types and terms can be inter-converted as follows:

- The curried type  $(\tau_1 \Rightarrow \dots \Rightarrow (\tau_n \Rightarrow \tau) \dots)$  can be written in uncurried form as  $\tau_1 \times \dots \times \tau_n \Rightarrow \tau$  and vice versa; and
- The curried term  $t(t_1) \dots (t_n)$  can be written in uncurried form as  $t(t_1, \dots, t_n)$  and vice versa.

### 3.2 Ontologies as Contexts

ORPs operate on higher-order ontologies that are formalized and implemented as contexts (Ballarin, 2004).

**Definition 5 (Higher-Order Ontologies)** *Let  $\mathbb{O}$  be the meta-type of ontologies. A higher-order ontology,  $O:\mathbb{O}$ , is a pair  $\langle S, A \rangle$ , where  $S$  is the signature of  $O$  and  $A$  is the set of axioms. We define  $\text{Sig}(O) ::= S$  and  $\text{Ax}(O) ::= A$ . The signature,  $S$ , is a set of type declarations for all the constants in  $C$ , i.e., the elements of  $S$  are of the form  $c:\tau$ , where  $c \in C$  and  $\tau \in \mathcal{T}$ . We will assume that a potentially infinite set of variables is provided for each type in  $\mathcal{T}$ . With this assumption, the signature defines a set of terms  $T$ . These terms are called the language of the ontology, denoted  $\text{Lang}(O)$ . The set of sentences of the ontology,  $\Phi$ , are the subset of terms that have type  $\text{bool}$  and no free variables, i.e.,  $\Phi ::= \{\phi \in T \mid \phi:\text{bool} \wedge FV(\phi) = \{\}\}$ . To reduce clutter, we will sometimes omit outermost universal quantifiers from sentences.*

An axiom,  $a \in A$ , is a sentence that is assumed to be true. When the axioms are closed under the rules of inference they define the set of theorems of the ontology, denoted  $\text{Th}(O)$ . If a sentence  $\phi$  is a theorem of ontology  $O$  we write  $O \vdash \phi$ .

An ontology  $O$  is said to be inconsistent if  $\perp$  is a theorem, i.e.,  $O \vdash \perp$ .

To reduce clutter we adopt the convention that  $\text{Sig}(O_i)$  is abbreviated to  $S_i$  and  $\text{Ax}(O_i)$  is abbreviated to  $A_i$ . Note that, depending on the context,  $S_i$  and  $A_i$  may refer to the local signatures and axioms, respectively.

$O \vdash \phi$  is an example of a *meta-assertion*, i.e., a sentence of a meta-ontology. A meta-ontology is a higher-order ontology whose domain is other ontologies, i.e., in which the constants denote object-level expressions, names of ontologies and models, etc. Ontology evolution consists of simultaneous inference in both object- and meta-ontologies. For our application, the meta-level is responsible for encoding the object-level logic and the ORPs. Object-logics are formed and manipulated by Isabelle's meta-logic, Isabelle/Pure, which serves as a logical framework for encoding arbitrary object-logics. However, the meta-logic in which the ORPs are expressed require more than just manipulating terms, but also generating proof obligations for diagnosis, gathering the preferred diagnosis, analysing dependencies between ontologies, and so forth. The meta-logic in



which the ORPs are expressed is encoded in the ML (Meta-Language) level, which interfaces with Pure.

An example of a polymorphic axiom inter-relating multiple polymorphic constants is:

$$\forall x, y, z: \tau. x +_{\tau} y =_{\tau} z \iff z -_{\tau} x =_{\tau} y$$

So, whatever the domain of application, we may insist on a fixed relationship between  $=_{\tau}$ ,  $+_{\tau}$  and  $-_{\tau}$ .

The rules of inference of this logic are the natural deduction rules provided by the Isabelle theorem prover for HOL and are as documented in Nipkow et al. (2002)[ch5]. The semantics of the logic are described in Gordon and Pitts (1994). This semantics defines the concept of *interpretation* of an ontology  $O$  as giving a meaning to each  $c: \tau \in \text{Sig}(O)$  and, hence, a truth value to each sentence  $\phi \in \text{Lang}(O)$ . We write  $M \models \phi$  if  $\phi$  is assigned  $\top$  by the interpretation  $M$ . If  $\forall \phi \in \text{Ax}(O). M \models \phi$  then  $M$  is said to be a *model* of  $O$ . We can regard the real world as defining an interpretation that is intended to be a model of all our ontologies. We will refer to this interpretation as the *preferred model*. Note that the very notion of reality is a deep philosophical issue. In our approach we adopt a deterministic view that responds to the needs of a logical representation. A logical model is fully determined by the truth value of the ground, atomic formulae. We can treat the real world as a model in the following sense. To discover the truth value of a ground, atomic formulae we conduct an experimental observation designed to discover its truth value. Of course, a change of representation will affect how such experiments are designed. For instance, if we believe that the value of a function does not depend on the day of the week then we will observe its presumed unique value on any convenient day. If we subsequently decide that the function *does* vary with the day, then we will need to take seven different observations. Thus the preferred model will subtly change as the ontology evolves, but will still correspond to the real world as currently understood at any given time.

**Definition 6 (Combining Ontologies)** *The combination,  $O_x \oplus O_y$ , of two ontologies,  $O_x$  and  $O_y$  is defined as:*

$$O_x \oplus O_y ::= \langle \text{Sig}(O_x) \cup \text{Sig}(O_y), \text{Ax}(O_x) \cup \text{Ax}(O_y) \rangle$$

*provided no clash of constant types arises, i.e.,*

$$\forall c \in C. c: \tau_x \in \text{Sig}(O_x) \wedge c: \tau_y \in \text{Sig}(O_y) \implies \tau_x = \tau_y$$

$\oplus : \mathbb{O} \times \mathbb{O} \Rightarrow \mathbb{O}$  *is a meta-constant.*

Note that the language of the combined ontologies contains the union of their languages and the axioms of the combined ontologies contains the union of their axioms.

The associativity and commutativity of  $\oplus$  follows trivially from the associativity and commutativity of  $\cup$ . We can, therefore, drop the parentheses from  $O_1 \oplus (O_2 \dots (O_{n-1} \oplus O_n) \dots)$  and write  $O_1 \oplus \dots \oplus O_n$  unambiguously.

**Definition 7 (Local vs Global (In)consistency)** *Consider the set of ontologies  $\{O_i: \mathbb{O} \mid 1 \leq i \leq n\}$ . The set is said to be locally (in)consistent if each  $O_i$  is (in)consistent. It is said to be globally (in)consistent if  $O_1 \oplus \dots \oplus O_n$  is (in)consistent.*

**Definition 8 (Ontology Fault)** *By an ontology fault we mean one of the following two cases:*

*Over-specification:* It is possible to prove a false theorem, i.e.,  $O \vdash \phi$  but  $\mathcal{M} \models \neg \phi$ , where  $\mathcal{M}$  is a preferred model of ontology  $O$  and  $\phi$  is a sentence in  $\text{Lang}(O)$ .

*A special case of over-specification is inconsistency, i.e.,  $O \vdash \perp$ , where  $\perp$  represents false, since  $\perp$  is false in all models. Note that inconsistency would arise if we tried to combine two ontologies, say  $O_x$  and  $O_y$ , where there exists a sentence  $\phi$  such that  $O_x \vdash \phi$  and  $O_y \vdash \neg \phi$ , since  $\perp$  would then be a theorem of the combined ontology  $O_x \oplus O_y$ .*

*Under-specification:* It is impossible to prove a true sentence, i.e.,  $O \not\vdash \phi$  but  $\mathcal{M} \models \phi$ , where  $\mathcal{M}$  is a preferred model of ontology  $O$  and  $\phi$  is a sentence in  $\text{Lang}(O)$ .

*A special case of under-specification is redundancy, i.e., that  $O \not\vdash t_1 = t_2$  but  $\mathcal{M} \models t_1 = t_2$ , where  $\mathcal{M}$  is a preferred model of ontology  $O$  and  $t_1$  and  $t_2$  are terms of the same type in  $\text{Lang}(O)$ .*

### 3.3 Multiple Ontologies

Most work in the area of ontological fault diagnosis attempts to detect a fault from a single ontology. Even though some work with networked ontologies (Ji et al., 2009), the network is typically collapsible into a single, large ontology. Our method for determining whether the given ontologies require repair focuses on the identification of the pattern describing an underlying fault across multiple ontologies, which is the purpose of the trigger formulae in ORPs. Effectively, a single ontology setup will render our proposed method of fault diagnosis useless. This means that a representation containing multiple ontologies are essential to performing precise diagnosis of ontological faults.

A goal in the implementation of GALILEO is to ease the restriction on the languages used by multiple ontologies, enabling a more flexible approach to modelling. It is fairly commonplace in dynamic, unregulated environments to deal with ontologies that exhibit different languages; for instance, agents fail to communicate with each other due to a lack of a shared understanding of the languages, which is caused by a heterogeneity of languages. If we regard each agent as having its own separate ontology, then such heterogeneity arises across multiple discrepant ontologies. To handle heterogeneous ontologies, we construct bridging ontologies.

**Definition 9 (Bridging Ontologies)** A bridge ontology,  $B_{O_1 \dots O_n} : \mathbb{O}$  between  $n$  ontologies,  $O_1, \dots, O_n$  is defined as:

$$B_{O_1 \dots O_n} ::= \langle \bigcup_{i=1}^n \text{Sig}(O_i), \{\Psi_j(\phi_1^j, \dots, \phi_n^j) \mid j \in [m]\} \rangle$$

provided no clash of constant types arises, i.e.,

$$\forall c \in C, \forall i, j \in [n]. c : \tau_i \in \text{Sig}(O_i) \wedge c : \tau_j \in \text{Sig}(O_j) \implies \tau_i = \tau_j$$

where  $\phi_i \in \text{Lang}(O_i)$  for  $i \in [n]$ , where  $[n] = \{i \mid 1 \leq i \leq n\}$ .

That is, the signature of the bridge ontology is the union of signatures of the  $n$  bridged ontologies; the set of axioms of the bridge ontology consists of  $m$  sentences that each asserts a relationship between formulae in the languages of the  $n$  bridged ontologies.

We can now combine  $B_{O_1 \dots O_n}$  with one or more of the  $O_i$  in order to make its theorems comparable to one or more of the remaining  $O_j$ . Note that combining  $B_{O_1 \dots O_n}$  with *all* of the  $O_i$ s will often create an inconsistent ontology.

Note that the notion of bridge used in this article as well as the axioms that are within a bridge resolve any linguistic differences between multiple ontologies. This notion is comparable to both the notions of lifting axioms and of bridging rules, which were proven equivalent in (Bouquet and Serafini, 2003).

### 3.4 Factorisation

In order to design a highly general mechanism for detecting faults in multiple ontologies, we must enable at least some inference to be performed across even globally inconsistent ontologies. For instance, given a predictive ontology containing definitions and a sensory ontology containing pure empirical data, a conflict is deducible only if the definitions are instantiated by the empirical data. Clearly, reasoning with the combination of these globally inconsistent ontologies is virtually meaningless, but the inconsistency itself may be circumvented by excluding certain axioms from the combination. Our solution is to enable ORPs to be performed not on the original input ontologies themselves, but on the ontologies in a *factorised representation* of the original.

**Definition 10 (Factorised Representation)** The factorised representation  $\mathcal{F}(O)$  of an ontology  $O : \mathbb{O}$  is defined as:

$$\mathcal{F}(O) ::= \{O' \mid O' \preceq O\}$$

where

$$\langle S', A' \rangle \preceq \langle S, A \rangle \iff S' = S \wedge A' \subseteq A \wedge A' \neq \emptyset.$$

A factorised representation of an ontology with  $n$  axioms is essentially a set of ontologies in which each ontology contains a unique  $k$ -combination of original axioms for  $1 \leq k \leq$

$n$ . The idea of working with a factorised representation of ontologies is loosely based on the concept of a factorised representation of a set of independent variables in Bayesian Networks (Jensen, 1996). Consequently, each ontology in a factorised representation is a sub-ontology of the ontology being factorised. The models satisfying any subset of the original axioms must also satisfy an ontology in the factorised representation and the models satisfying any ontology in the factorised representation must also satisfy a subset of the original axioms. Sub-ontologies have been proven useful and adopted to solve specific reasoning tasks about ontologies (Haase et al., 2005; Du et al., 2008). We will use the superscript  $'$  and multiples thereof to denote the ontologies in a factorised representation, e.g.,  $O', O''$ , i.e. are sub-ontologies of the original ontology,  $O$ .

Note the following rather trivial facts, which allow us to describe below the relationship between the preferred model of an ontology and the models the ontology's factorised representation.

**Proposition 1** Given an ontology  $O : \mathbb{O}$ , the following holds:

$$\forall \emptyset \neq A' \subseteq \text{Ax}(O). \exists O' \in \mathcal{F}(O). \text{Ax}(O') = A'.$$

**Proposition 2** Given an ontology  $O : \mathbb{O}$ , the following holds:

$$\forall O' \in \mathcal{F}(O). \exists A' \subseteq \text{Ax}(O). \text{Ax}(O') = A'.$$

From Propositions 1 and 2, we claim that the preferred model of an ontology, which itself should be a model of at least one of the original axioms, is a model of at least one of the ontologies in the corresponding factorised representation. This is because the factorised representation must contain all possible sub-ontologies of the corresponding ontology. Working with a factorised representation, therefore, essentially enables reasoning to be performed over any subset of the axioms of the input ontologies. In the case where the environment contains multiple factorised representations, each corresponding to a different input ontology, an ontology in one factorised representation could be combined with one in another. In effect, a subset of the axioms of an input ontology can now be combined with the whole of another input ontology. This is a powerful approach and increases the generality of the repair mechanism, because it enables some reasoning across globally inconsistent ontologies – the reasoning can be focussed on only those ontologies in the factorised representations that do *not* contain the axioms that would otherwise induce global inconsistency. The combination operator  $\oplus$  enables the reuse of ontologies for building factorised representations as networks, so that the number of ontologies created can be kept to a minimum.

### 3.5 Ontology Repair Plans

The general notion of ORP in terms of the logic defined above is the following.

**Definition 11 (Ontology Repair Plans)** *An ontology repair plan is a pair  $\langle \text{Trigger}, \text{Repair} \rangle$ , where:*

*Trigger: is a set of meta assertions of the form  $O \vdash \phi$  or  $O \not\vdash \phi$  which collectively create an ontology fault.*

*Repair: is a set of repair operations of the form  $v(O) ::= \pi(O)$ , where  $v(O)$  denotes a repaired ontology and  $\pi$  generically is a generic operation on the signature and/or axioms of the original ontology  $O$ . If each  $O$  is replaced by  $v(O)$  in Trigger to form  $v(\text{Trigger})$  then  $v(\text{Trigger})$  no longer creates the ontology fault detected in Trigger<sup>4</sup>.*

The signature of the repaired ontology,  $\text{Sig}(v(O))$ , may contain type declarations for constants that are not declared in the signature of the original ontology,  $\text{Sig}(O)$ , and vice versa. One consequence is that a model of  $O$  may not be even be an interpretation of  $v(O)$  and vice versa. This makes it difficult to give a model-based semantics to the operations of ontology repair plans. Note that this problem is caused by the inclusion of signature change in ontology evolution and not by the use of higher-order logic. We are exploring other possibilities, but assigning a semantics to our version of higher-order ontology evolution remains further work.

In the following sections we discuss three ORPs, their formalization, implementation and testing. The ORP called *Where's My Stuff?* (WMS) is triggered in situations where a theoretical prediction in  $O_x$  conflicts with sensory information derived from experiments in  $O_y$  possibly through a bridging ontology  $B$ . In the case where bridging is required,  $O_y$  can be instantiated to a combination of  $B$  and the sensory ontology in consideration using the  $\oplus$  operator. WMS deploys an addition-strategy that is quite common in Physics. For instance, in order to account for unpredictable yet observed gravitational behaviours in the orbit of a planet or in the stellar orbital velocity in a galaxy, astronomers postulate the presence of an additional unobserved planet or, resp., of dark matter. Accordingly, WMS redefines the contradictory function (in the examples, the functions orbit, resp., orbital velocity) as the sum of a *visible* part (i.e. the amount calculated by the original function) and an *invisible* part (i.e. the amount that can only indirectly be observed).

The ORP called *Inconstancy* is triggered in situations where sensory information is derived from experiments run under different circumstances  $O_{y,1} \dots O_{y,n}$ . If such experiments present variations (e.g. variations in temperature) in the value of a function *stuff* which, according to  $O_x$ , should be constant (e.g. the volume to pressure ratio of a gas), *Inconstancy* repairs the theoretical ontology by changing the

signature and making the constant quantity dependent on the varying quantity.

The ORP called *Unite* is triggered in situations where the observations in the sensory ontology  $O_y$  for two possibly distinct entities in the theoretical ontology  $O_x$  fully match, and this is grounds for identifying the two functions as one. *Unite* is the converse of WMS: it is not triggered by a contradiction, but by an equality of two *stuff*'s relative to a chosen defining property; the repair is to equate in  $O_x$  the two *stuff*'s, so as to eliminate the redundancy.

Other ORPs perform other types of repairs, such as unifying two theoretical ontologies; changing the type structure of a function to let it fit a given dataset; drawing an analogy between two theoretical ontologies; skolemizing a theorem containing an existential quantifier and transforming a set of unary predicates into a unary function. Some of these ORPs are discussed as part of further work in Section 9.

Finally, as further discussed in Section 8.3, some of the diagnoses found by GALILEO suggest that experimental rather than predictive error is the cause of the fault. Therefore, although the accent of the ORPs discussed below is on repairing the predictive elements of the ontologies, a repair may principally work on the experimental ontology, corresponding to a faulty experiment, rather than on the theoretical one.

## 4 The *Where's My Stuff?* ORP

The *Where's My Stuff?* ORP, described in Figure 3, is triggered when the predicted value returned by a function conflicts with the observed value of the same function. It assumes two ontologies: an ontology  $O_x$  representing the current state of a predictive Physics theory and a heterogeneous ontology  $O_y$  representing some sensory information arising from an experiment. If the ontologies are heterogeneous, the reasoning performed may possibly involve a bridging ontology mapping between heterogeneous formulae across ontologies. Suppose the function  $f$  measures some property of *stuff*. There are two possible ways to identify the conflict: with equation (15), where  $f(\text{stuff})$  is equal or less than a particular value in  $O_y$ , or with equation (16), where  $f(\text{stuff})$  is equal or greater than a particular value in  $O_x$ . The repair is to split *stuff* into three: *visible stuff*, *invisible stuff*, and *total stuff*, defining *invisible stuff* in terms of total and *visible stuff*'s in the repaired  $O_x$ ,  $v(O_x)$ , as in equations (17, 20). The new  $O_y$ ,  $v(O_y)$ , is the same as  $O_y$  except for the renaming of *stuff* to *stuff<sub>vis</sub>*, as in equation (21).

### 4.1 WMS's application to the discovery of latent heat

Until the second half of the 18th century, the chemical/physical notion of heat was conflated with the notion of temperature and it was seen as a function of a temporal quantity, e.g.

<sup>4</sup> Although, other ontology faults may still remain to be detected and repaired, as well faults newly introduced by the repair plan.

Trigger:  $O_x$  is the predictive ontology and  $O_y$  is the observational one. If  $f(stuff)$  has two different values in  $O_x$  and  $O_y$ , then the following formula will be triggered:

$$\exists O_x, O_y: \mathbb{O}, \tau, \tau': Types, f: \tau \Rightarrow \tau', stuff: \tau, v: \tau'. \\ (O_x \vdash f(stuff) >_{\tau'} v \wedge O_y \vdash f(stuff) \leq_{\tau'} v) \vee \\ (O_x \vdash f(stuff) \geq_{\tau'} v \wedge O_y \vdash f(stuff) <_{\tau'} v) \quad (15)$$

where  $O \vdash \phi$  means that formula  $\phi$  is a theorem of ontology  $O$ ;  $t: Types$  means  $t$  is a type;  $o: \mathbb{O}$  means  $o$  is an ontology;  $>_{\tau'}$  is the greater-than operator for  $\tau'$ .

Repair: Two new kinds of stuff and a definition of the *invisible* stuff are introduced.

$$stuff \sigma_{invis} ::= stuff -_{\tau} stuff \sigma_{vis} \quad (17)$$

When  $stuff$  is a constant, the substitution  $\sigma_{vis}$  just replaces it with a new constant standing for the visible stuff; when  $stuff$  is compound, the replacement is more complex, but still automatable. Similar remarks hold for  $\sigma_{invis}$ .

Let  $v(O_x)$  and  $v(O_y)$  be the repaired ontologies. The signatures for the new ontologies are updated in terms of those of the old as follows:

$$Sig(v(O_x)) ::= \{stuff_{vis}: \tau, stuff_{invis}: \tau\} \cup Sig(O_x) \quad (18)$$

$$Sig(v(O_y)) ::= \{stuff_{vis}: \tau\} \cup Sig(O_y) \setminus \{stuff: \tau\} \quad (19)$$

The axioms for the new ontologies are updated in terms of those of the old as follows:

$$Ax(v(O_x)) ::= \{stuff \sigma_{invis} ::= stuff -_{\tau} stuff \sigma_{vis}\} \cup Ax(O_x) \quad (20)$$

$$Ax(v(O_y)) ::= \{\phi \{stuff / stuff \sigma_{vis}\} \mid \phi \in Ax(O_y)\} \quad (21)$$

To effect the repair, the axioms of  $v(O_x)$  are the same as those of  $O_x$  except for the addition of the new definition; the axioms of  $v(O_y)$  are the same as those of  $O_y$  except for the renaming of the original stuff to the visible stuff. Note that in case the value of  $f(stuff)$  in  $O_x$  is smaller than that in  $O_y$ ,  $O_y$  is repaired as  $O_x$  in (20) and  $O_x$  as  $O_y$  in (21). Because  $O_x$  and  $O_y$  can be instantiated to a combination of ontologies, if the instantiation of  $O_x$  is a combination of ontologies and some of these ontologies are used in the instantiation of  $O_y$ , then they are excluded from being repaired by (21).

Fig. 3: *The Where's My Stuff?* ORP

flow (Wiser and Carey, 1983). Flow was the cause of the change of temperature of a physical body put in direct contact with another physical body at a different temperature. This pre-modern view of heat and temperature can be rationally reconstructed in the following equation:

$$\Delta T \equiv Q = m \times \Delta t \quad (22)$$

where  $Q$  is the heat absorbed or released by a body and is equivalent to  $\Delta T$ , the body's difference in temperature at the start and at the end of the flow,  $m$  is the body's mass,  $\Delta t$  is the flow of heat, measured over time, from the hotter to the cooler body.

In the period 1759-1763 the discoveries of specific and of latent heat by Joseph Black brought a revision of the view underlying equation (22) and established the modern view, based on a distinction between the quantity of heat transferred to a physical body and the rise in temperature undergone by the body. On the one hand, the notion of specific

heat capacity (1759) accounted for the fact that when equal masses of different materials at equal temperatures absorb the same quantity of heat (i.e. when exposed to equal flows) they undergo different rises in temperature. This is due to a material specific capacity constant  $c$ . On the other hand, the case of single-substance bodies led to the formulation of the theory of latent heat (1761). A melting block of, for instance, ice releases heat at constant temperature. Such heat is proportional to a constant  $L$  which is specific to the body's material as well as to the phase-transition. This modern view of heat and temperature is represented by the following equation:

$$Q = m \times \Delta T \times c + m \times L \quad (23)$$

where  $Q$  is the heat put into or taken out of the body,  $m$  is the mass of the body,  $c$  is the specific heat capacity of the body's substance,  $\Delta T$  is the change in temperature,  $L$  is the specific latent heat of the substance during the considered phase-transition.

Note that the evolution from equation (22) to equation (23) marks a shift in the meaning of the variables:  $Q$  is not measured solely by the temperature difference of the body,  $\Delta T$  (which now has its own place in the equation). Also, there is no need any more to make explicit reference to a temporal notion such as flow ( $\Delta t$ ).

Figure 4 represents the state of the theory of heat according to equation (22) and to the experimental results that led to the introduction of latent heat. At the top,  $S_{1,2}$  contains the shared signature of the two ontologies of Physics,  $O_1$  and  $O_2$ . For instance, *Melting* is defined as a subtype of event (*Evt*), *Start* as a function from *Evt* to the type of the real numbers  $\mathbb{R}$ . Note that in this article, wherever appropriate, models of Physics laws and of experiments make use of an event-based representation, which makes them compatible with the Event Calculus (EC) (Kowalski and Sergot, 1986).  $O_1$ , whose signature is  $S_1$ , contains new signature elements such as the duration of an event (*Duration*) and represents equation (22) by equation (26). It also contains value assertions for *Mass* and *Duration* that allow this ontology to predict that the quantity of *Heat* released by  $H_2O$  during the *Melting* event is greater than zero. On the other hand,  $O_2$ , whose signature is  $S_2$ , represents the experimental setup: it contains other signature elements for the instruments (such as *ThmIn* for the thermometer used in the measurements) and it defines the difference in temperature of a body (*TempDiff*) as the difference in height (*Hght*) of the mercury column (*MerOf*) of a thermometer stuck in the body throughout an event. It also contains value assertions for the relevant observations of the height of the mercury.

Bridge  $B_{O_1', O_2'}$  aligns heterogeneous formulae between sub-ontologies  $O_1'$  and  $O_2'$ ; for instance, formulae containing *Heat* in  $O_1'$  and formulae containing *TempDiff* in  $O_2'$ . This allows the contradiction with  $O_1$  to become explicit to match

$$S_{1,2} ::= \{ \text{Mass:Obj} \Rightarrow \text{Evt} \Rightarrow \mathbb{R}, \text{Melting:Evt}, \text{Start:Evt} \Rightarrow \mathbb{R}, \dots \} \quad (24)$$

$$S_1 ::= \{ \text{Duration:Evt} \Rightarrow \text{Time}, \text{Heat:Obj} \Rightarrow \text{Evt} \Rightarrow \mathbb{R}, \text{HeatDiff:Obj} \Rightarrow \text{Evt} \Rightarrow \mathbb{R} \} \quad (25)$$

$$A_1 ::= \{ \forall o:\text{Obj}, e:\text{Evt}. \text{Heat}(o, e) = \text{Mass}(o, e) \times \text{Duration}(e), \quad (26)$$

$$\text{Mass}(\text{H}_2\text{O}, \text{Melting}) > 0, \quad (27)$$

$$\text{Duration}(\text{Melting}) > 0 \} \quad (28)$$

$$S_2 ::= \{ \text{Hght:Obj} \Rightarrow \text{Time} \Rightarrow \mathbb{R}, \text{TempDiff:Obj} \Rightarrow \text{Evt} \Rightarrow \mathbb{R} \} \quad (29)$$

$$A_2 ::= \{ \forall o:\text{Obj}, e:\text{Evt}. \text{TempDiff}(o, e) = \text{Hght}(\text{MerOf}(\text{ThmIn}(o, \text{Start}(e)))) - \text{Hght}(\text{MerOf}(\text{ThmIn}(o, \text{End}(e))))), \quad (30)$$

$$\text{Hght}(\text{MerOf}(\text{ThmIn}(\text{H}_2\text{O}, \text{Start}(\text{Melting})))) = 5, \quad (31)$$

$$\text{Hght}(\text{MerOf}(\text{ThmIn}(\text{H}_2\text{O}, \text{End}(\text{Melting})))) = 5 \} \quad (32)$$

$$\text{Sig}(B_{O_1', O_2'}) ::= S_1 \cup S_2 \cup S_{1,2} \quad (33)$$

$$\text{Ax}(B_{O_1', O_2'}) ::= \{ \forall o, e. \text{Mass}(o, e) > 0 \implies \text{Heat}(o, e) = \text{TempDiff}(o, e) \} \quad (34)$$

Fig. 4: *Initial model of the Latent Heat case study.*  $S_i$  denotes the local signature of  $O_i$  and  $A_i$  denotes the set of local axioms of  $O_i$ . Ontologies  $O_1'$  and  $O_2'$  are two sub-ontologies of  $O_1$  and  $O_2$ , respectively.  $O_1'$  contains only the axiom (26) and  $O_2'$  contains all axioms of  $O_2$ . Bridge  $B_{O_1', O_2'}$  merges  $O_1'$ 's and  $O_2'$ 's signatures, which are the same as those of  $O_1$  and  $O_2$ , and maps their concepts, for instance, *Heat*, the amount of heat absorbed or released by a body, is mapped on *TempDiff*, the temperature difference undergone by the body. This allows the contradiction between  $O_1$  and  $O_2$  via the bridge  $B_{O_1', O_2'}$  about the value of *Heat* to become explicit to match the WMS trigger formula, because during a *Melting* event there is no change in temperature, although  $O_1$  predicts it.

the WMS trigger formula (15):

$$O_1'' \vdash \text{Heat}(\text{H}_2\text{O}, \text{Melting}) > 0 \quad (35)$$

$$O_2' \oplus B_{O_1', O_2'} \oplus O_1' \vdash \text{Heat}(\text{H}_2\text{O}, \text{Melting}) = 0 \quad (36)$$

where  $O_1''$  is a sub-ontology in the factorised representation of  $O_1$  containing all of the original axioms such that the above holds. Table 1 outlines the sub-ontologies generated by the factorisation of  $O_1$ . Only with  $O_1'$  the value of  $\text{Heat}(\text{H}_2\text{O}, \text{Melting})$  can be inferred using the empirical data

$O_1'$	{(26)}
$O_1''$	{(26), (27), (28)}
$O_1'''$	{(26), (27)}
$O_1''''$	{(26), (28)}
$O_1'''''$	{(27), (28)}
$O_1''''''$	{(27)}
$O_1'''''''$	{(28)}

Table 1: Axioms of all sub-ontologies generated by the factorisation of  $O_1$ .

and the bridge; the combination with  $O_1''$  will induce inconsistency and the others do not contain the relevant axioms to allow the inference.

Given the substitution:

$$\{ O_1''/O_x, B_{O_1', O_2'} \oplus O_2'/O_y, \lambda x.x(\text{H}_2\text{O}, \text{Melting})/f, \text{Heat}/\text{stuff}, 0/v \}$$

WMS repairs the two ontologies as in  $v(O_x)$  (i.e. by adding to  $O_1''$  the distinction between visible and invisible *Heat*) and as in  $v(B_{O_1', O_2'})$  (i.e. by renaming in  $B_{O_1', O_2'}$  all occurrences of *Heat* to *Heat<sub>vis</sub>*). Note that there are no occurrences of *Heat* in  $O_2'$  so it doesn't receive repair.

$$\text{Ax}(v(O_1'')) ::= \{ \text{Heat}_{\text{invis}} ::= \text{Heat} - \text{Heat}_{\text{vis}} \} \cup \text{Ax}(O_1'') \quad (37)$$

$$\text{Ax}(v(B_{O_1', O_2'})) ::= \{ \phi \{ \text{Heat}/\text{Heat}_{\text{vis}} \} \mid \phi \in \text{Ax}(B_{O_1', O_2'}) \} \quad (38)$$

The effect of repair is not limited to  $v(O_1'')$  but is propagated to the entire corresponding factorised representation. Every sub-ontology in the representation receives the same treatment of repair, i.e.:

$$\forall O \in \mathcal{F}(O_1).$$

$$\text{Ax}(v(O)) ::= \text{Heat}_{\text{invis}} ::= \text{Heat} - \text{Heat}_{\text{vis}} \} \cup \text{Ax}(O)$$

$$\forall O \in \mathcal{F}(B_{O_1', O_2'}).$$

$$\text{Ax}(v(O)) ::= \{ \phi \{ \text{Heat}/\text{Heat}_{\text{vis}} \} \mid \phi \in \text{Ax}(O) \}.$$

## 4.2 WMS's application to the postulation of dark matter

The theoretical existence of dark matter is based on various sources of evidence, including the rotational velocities of stars in spiral galaxies, which exceeds the predicted orbital velocities<sup>5</sup>, as first observed by Rubin (Rubin et al., 1980).

<sup>5</sup> It is assumed that *observed* orbital velocity and rotational velocity for stars are the same. Observations were initially based on gas clouds (HII regions) rather than stars, and only later were the observations based on both gases and stars. We simplify the Physics by considering only stars.

Given the observed distribution of mass in these galaxies, Newtonian dynamics predicts that orbital velocities decrease inversely with the square root of the distance from the galactic centre, or the *radius* (equation (42)). However, the observed velocity was almost constant out to large radii. Rubin's conclusion was that some invisible matter exerts a gravitational force on these stars, causing the unexpectedly high orbital velocities. The theoretical notions that play a role in this case are Newton's second law of motion (equation (39)), the law of gravitational attraction in circular orbits (equation (40)) and the law of centripetal acceleration (equation (41)), which combined allow to derive the orbital velocity of a body (equation (42)) at distance  $r$  from the body at the centre of the orbit.

$$F = ma \quad (39)$$

$$F = \frac{GMm}{r^2} \quad (40)$$

$$a = \frac{v_{orb}^2}{r} \quad (41)$$

$$v_{orb} = \sqrt{\frac{G \times M}{r}} \quad (42)$$

where  $F$  is the force applied to an orbiting body,  $m$  is the mass of the body,  $a$  is its acceleration,  $M$  is the mass of the body at the centre of the orbit,  $G$  is the gravitational constant,  $r$  is the radius, or distance between the orbiting body and the body at the centre of the orbit,  $v_{orb}$  is the orbital velocity.

On the experimental side, orbital velocity is calculated as rotational velocity, based on spectrographic data and using, among others, equations (43, 44, 45).

$$z = \frac{\lambda - \lambda_0}{\lambda_0} \quad (43)$$

$$v_{rad} = c \times z \quad (44)$$

$$v_{rot} = \frac{v_{rad} - v_{sys}}{\sin(i)} \quad (45)$$

where  $z$  is the redshift of a radiation  $\lambda$  with respect to a reference  $\lambda_0$ ,  $c$  is the speed of light,  $v_{rad}$  is the radial velocity of a body (e.g. a star) along the line of observation,  $v_{sys}$  is the velocity of the system (e.g. a galaxy) to which the observed body belongs,  $v_{rot}$  is the rotational velocity of the object.

Figure 5 represents the state of the theory of galactic orbital velocity according to equation (42) and to the observations yielded by equations (43) to (45), which led to the postulation of dark matter. At the top,  $S_{3,4}$  contains the shared signature where, for instance, the type *Rad* for distances from the galactic centre is defined as a function that maps the product of objects (e.g. stars) and object sets (e.g. galaxies) and an observation event onto the type *Dst* for distances (this in turn maps onto real numbers). Note that  $S_{3,4}$  is assumed to inherit, from some higher mathematical ontologies, knowledge about shapes of curves, e.g., *UpFlat*

$$S_{3,4} ::= \{UpFlat: \mathbb{R} \Rightarrow \mathbb{R}, Rad: (Obj * Obj Set) \Rightarrow Evt \Rightarrow Dst \dots\} \quad (46)$$

$$S_3 ::= \{GphA: Gly \Rightarrow Evt \Rightarrow \mathbb{R}, OV: Gly \Rightarrow Evt \Rightarrow \mathbb{R}, \dots\} \quad (47)$$

$$A_3 ::= \{\forall e: Evt, g: Gly, r: Dst. OV(g, e) = \sqrt{\frac{G \times \sum s \in g. Mass(s, e)}{r}}, \quad (48)$$

$$\forall e: Evt. GphA(Glxy71, e) = OV(Glxy71, e), \quad (49)$$

$$G = 6.673 \times 10^{-11}, \quad (50)$$

$$\sum s \in Glxy71_{Rad(Star1, Glxy71)}. Mass(s, Obs1) = 100, \quad (51)$$

$$\sum s \in Glxy71_{Rad(Star9, Glxy71)}. Mass(s, Obs9) = 110\} \quad (52)$$

$$S_4 ::= \{GphB: Gly \Rightarrow Evt \Rightarrow \mathbb{R}, RtV: Str \Rightarrow Gly \Rightarrow Evt \Rightarrow \mathbb{R}, \dots\} \quad (53)$$

$$A_4 ::= \{\forall e: Evt, s: Str, g: Gly. RdV(s, g, e) = c \times RedShift(s, g, e), \quad (54)$$

$$\forall e: Evt, s: Str, g: Gly. RtV(s, g, e) = \frac{RdV(s, g, e) - SyV(g, e)}{\sin(Inc(g, e))}, \quad (55)$$

$\forall s: Str.$

$$GphB(Glxy71, Rad(s, Glxy71)) = RtV(s, Glxy71, Obs9), \quad (56)$$

$$c = 299792458, \quad (57)$$

$$RedShift(Star1, Glxy71, Obs1) = 300, \quad (58)$$

$$RedShift(Star9, Glxy71, Obs9) = 300\} \quad (59)$$

$$Sig(B_{O_3', O_4'}) ::= \{S_3 \cup S_4 \cup S_{3,4}\} \quad (60)$$

$$Ax(B_{O_3', O_4'}) ::= \{\forall e: Evt, s: Str, g: Gly. RtV(s, g, e) = OV(g, Rad(s, g), e)\} \quad (61)$$

Fig. 5: *Initial model of the Dark Matter case study.* Ontologies  $O_3'$  and  $O_4'$  are in essence two sub-ontologies of  $O_3$  and  $O_4$ , respectively.  $O_3'$  contains only the axioms (48 - 50) and  $O_4'$  contains all the axioms of  $O_4$ .  $B_{O_3', O_4'}$  combines the signatures of  $O_3'$  and  $O_4'$ , which are the same as those of  $O_3$  and  $O_4$ , and aligns *RtV* with *OV*. This allows the contradiction with  $O_4$  to become explicit to match the WMS trigger formula (16). Note how the galaxy, its velocities and the graphs representing it are all modelled as higher order predicates in equations (48, 49, 54, 55, 56).

is the curve that has a positive gradient between zero and some point and a zero gradient thereafter. Such knowledge makes it possible to compare the predictions and the observations.  $O_3$  introduces the galaxy rotation curve *GphA* in its signature,  $S_3$ , a Keplerian curve computed according to equation (48) (which models equation (42)) where *Evt*, *Gly*, *Dst*, and *Str* denote types for representing events, galaxies,

distances, and stars, resp.;  $G$  denotes the universal gravitational constant;  $OV$ , the orbital velocity;  $g_r$ , the set of stars in the galaxy  $g$  up to distance  $r$ ;  $Mass$ , the mass of a body.  $O_4$  introduces  $GphB$  in its signature,  $S_4$ , the curve based on observations, computed according to equations (54) to (56) (which model equations (43) to (45)) where  $RtV$  denotes the rotational velocity;  $RdV$ , the radial velocity;  $SyV$ , the velocity of the galactic system relative to the observer;  $Inc$ , the inclination of the galaxy;  $RedShift$ , the shift in wave length;  $Glxy71$ , the galaxy being observed;  $Star1$  to  $Star9$ , stars in the observed galaxy; and,  $Obs9$ , the observation event.

Some symbols in  $O_4$ , such as  $RtV$  and  $RdV$ , are not in the language of  $O_3$ .  $B_{O_3',O_4'}$  links together the seemingly disparate terms by relating  $RtV$  in  $O_4$  to the  $OV$  in  $O_3$ . This allows the contradiction with  $O_4$  to become explicit to match the WMS trigger formula (16):

$$O_3'' \vdash GphA = UpFlat \quad (62)$$

$$B_{O_3',O_4'} \oplus O_4' \vdash GphA < UpFlat \quad (63)$$

where  $O_3''$  is a sub-ontology in the factorised representation of  $O_3$  containing all of the original axioms such that the above holds. Given the substitution:

$$\{O_3''/O_x, B_{O_3',O_4'} \oplus O_4'/O_y, \\ GphA/f, Glxy71/stuff, UpFlat/v\}$$

WMS repairs the two ontologies as in  $v(O_x)$  (i.e. by renaming in  $O_3'$  all occurrences of  $Glxy71$  to  $Glxy71_{vis}$ ) and as in  $v(O_y)$  (i.e. by adding to  $O_4''$  the distinction between visible and invisible  $Glxy71$ ).

$$Ax(v(O_3'')) ::= \{\phi\{Glxy71/Glxy71_{vis}\} \mid \phi \in Ax(O_3'')\} \quad (64)$$

$$Ax(v(O_4')) ::= \{Glxy71_{invis} ::= Glxy71 - Glxy71_{vis}\} \\ \cup Ax(O_4') \quad (65)$$

The effect of the repair is not limited to  $v(O_3'')$  and  $v(O_4')$  but can be propagated to the entire corresponding representation. Every sub-ontology in each representation receives the same treatment of repair, i.e.

$$\forall O \in \mathcal{F}(O_3).$$

$$Ax(v(O)) ::= \{\phi\{Glxy71/Glxy71_{vis}\} \mid \phi \in Ax(O)\}$$

$$\forall O \in \mathcal{F}(O_4).$$

$$Ax(v(O)) ::= \{Glxy71_{invis} ::= Glxy71 - Glxy71_{vis}\} \cup$$

$$Ax(O).$$

## 5 The Inconstancy ORP

The *Inconstancy* ORP, described in Figure 6, is triggered when, given an ontology  $O_x$  representing the current state of a physical theory and some ontologies  $O_{y,i}$  representing

Trigger:  $O_x$  is the predictive ontology,  $O_{y,i}(v(\mathbf{b}_i) =_{\tau'} v_i) \forall i \in [1, n]$  are the observational ones, made under the conditions  $v(\mathbf{b}_i) =_{\tau'} v_i$ . If  $f(stuff)$  is measured to take different values in at least two of these observational ontologies, then an unexpected variation is detected. The inference of the conflict w.r.t the observational data can be aided by ontologies  $O_{z,i}$ . We can formalise the trigger formulae as follows:

$$\exists O_x, O_{y,1}, \dots, O_{y,n}, O_{z,1}, \dots, O_{z,n}: \mathbb{O}, \\ \tau, \tau', \tau'', \tau''': Types, stuff: \tau, f: \tau \Rightarrow \tau', c, c_1, \dots, c_n: \tau', \\ v_1, \dots, v_n: \tau'', \mathbf{b}_1, \dots, \mathbf{b}_n: \tau'', v: \tau'' \Rightarrow \tau'' \\ (O_x \vdash f(stuff) =_{\tau'} c \wedge \quad (66)$$

$$O_{y,1}(v(\mathbf{b}_1) =_{\tau''} v_1) \oplus O_{z,1} \vdash f(stuff) =_{\tau'} c_1 \wedge \\ \vdots \\ O_{y,n}(v(\mathbf{b}_n) =_{\tau''} v_n) \oplus O_{z,n} \vdash f(stuff) =_{\tau'} c_n \wedge \quad (67)$$

$$\exists i, j \leq n. O_x \vdash c_i \neq_{\tau'} c_j \vee c_i \neq_{\tau'} c \quad (68)$$

ontology;

Repair: The repair is to change the signature of all the ontologies to relate the inconstancy,  $stuff$ , to the variad,  $v(\mathbf{y})$  via a new function  $F$  the value of which can be determined by regression analysis against the data from the sensory ontologies:

$$v(stuff) ::= \lambda \mathbf{y}. F(c, v(\mathbf{y})) \quad (69)$$

The signatures of the new ontologies are updated in terms of those of the old as follows:

$$Sig(v(O_x)) ::= \{F: \tau' \Rightarrow \tau'' \Rightarrow \tau', v(stuff): \tau'' \Rightarrow \tau\} \cup \\ Sig(O_x) \setminus \{stuff: \tau\} \quad (70)$$

$$Sig(v(O_{y,i}(v(\mathbf{b}_i) =_{\tau''} v_i))) ::= \{v(stuff): \tau'' \Rightarrow \tau\} \cup \\ Sig(O_{y,i}(v(\mathbf{b}_i) =_{\tau''} v_i)) \setminus \{stuff: \tau\} \quad (71)$$

$$Sig(v(O_{z,i})) ::= \{v(stuff): \tau'' \Rightarrow \tau\} \cup \\ Sig(O_{z,i}) \setminus \{stuff: \tau\} \quad (72)$$

The axioms of the new ontologies are calculated in terms of those of the old, where  $Ax(O)$  returns all axioms of  $O$ , as follows:

$$Ax(v(O_x)) ::= \\ \{\phi\{stuff/v(stuff)(\mathbf{y})\} \mid \phi \in Ax(O_x)\} \setminus \{stuff ::= c\} \cup \\ \{v(stuff) ::= \lambda \mathbf{y}. F(c, v(\mathbf{y}))\} \quad (73)$$

$$Ax(v(O_{y,i}(v(\mathbf{b}_i) =_{\tau''} v_i))) ::= \\ \{\phi\{stuff/v(stuff)(\mathbf{b}_i)\} \mid \phi \in Ax(O_{y,i}(v(\mathbf{b}_i) =_{\tau''} v_i))\} \quad (74)$$

$$Ax(v(O_{z,i})) ::= \\ \{\phi\{stuff/v(stuff)(\mathbf{y})\} \mid \phi \in Ax(O_{z,i})\} \quad (75)$$

where, each of the instantiations of the  $O_{y,i}$ s does not share ontologies if it is a combination of several ontologies; and, if the instantiation of  $O_{z,i}$  is a combination of ontologies and some of these make up the instantiation of  $O_x$ , then those ontologies do not receive the repair specified in (75).

Fig. 6: The Inconstancy ORP.

sensory information arising from experiments, the sensory ontologies give distinct values for function  $stuff$  in different circumstances. Suppose function  $v(\mathbf{b}_i)$  of the  $i^{th}$  sensory ontology, where vector  $\mathbf{b}_i$  contains variables distinguishing among these circumstances, returns distinct values in each of these circumstances, but is *not* one of the parameters in  $stuff$ , which does not depend on  $v(\mathbf{b}_i)$ . We call  $stuff$  the *inconstancy* and  $v(\mathbf{b}_i)$  the *variad*. The Inconstancy repair plan

establishes a relationship  $F$  between the variad and the inconstancy.

To discover the meaning of the function  $F$ , *Inconstancy* follows the tradition of Langley's BACON program (Langley et al., 1983). The ontologies  $O_{y,i}(v(\mathbf{b}_i) =_{\tau'''} v_i \dots)$  provide a useful collection of equations:  $F(c, v(\mathbf{b}_i)) =_{\tau'} c_i$  for  $i = 1, \dots, n$ . Regression analysis can be applied to these equations to approximate a definition of  $F$ . This hypothesis can then be tested by creating observations  $O_{y,k}(v(\mathbf{b}_k) =_{\tau'''} v_k \dots)$ , for new values of  $v(\mathbf{b}_k)$ , and confirming or refuting the hypothesis. If the inference relies on the use of a bridging ontology between a sensory and a theoretical ontology, then all occurrences of the inconstancy in the bridge are replaced by the new term with a free variable as its argument.

### 5.1 *Inconstancy* application to the MOND approach

The MODified Newtonian Dynamics (MOND), proposed by Milgrom in 1981, is alternative to dark matter and exemplifies how the same observations can trigger a different repair plan, *Inconstancy*. MOND suggests that the constant  $G$  is not a constant, but depends on the relative acceleration between the objects on which it acts. It is constant until the acceleration becomes very small and then it depends on this acceleration. So, the gravitational constant  $G$  (the inconstancy) in equation (42) can be repaired by redefining it as a function of itself and of the centripetal acceleration of stars due to the gravitational attraction between the star and the galaxy in which it belongs (the variad). Figure 7 extends and modifies the axiomatisation presented in Figure 5<sup>6</sup>.  $S_{5,6}$  extends  $S_{3,4}$  to include the symbol for centripetal acceleration  $Accel(s, e)$ . This is defined in  $O_5$  as a function  $F$  of orbital velocity and in  $O_6$  as a function of rotational velocity.  $O_{6,1}$  and  $O_{6,9}$  contain the data that allow to calculate according to equation (55) the rotational velocity for *Star1* to *Star9*, which has the same value. Therefore, according to equation (80) their acceleration varies with their radius. The bridges  $B_{O_{6,1}', O_5'}$  and  $B_{O_{6,9}', O_5'}$ , which are the same as  $B_{O_3', O_4'}$ , equate orbital and rotational velocity, allowing to calculate the local value for  $G$ , which varies as the variad.  $O_5'$  is a sub-ontology, containing the minimal number of axioms of  $O_5$  for the contradiction with  $O_5$  to become explicit by matching *Inconstancy's* trigger formulae (66) through (68), where  $G$  is calculated in each  $O_{6,i'}$  combined with  $O_5'$  via the bridge  $B_{O_{6,i'}, O_5'}$  by the inverse of equation (48) (the value of which is equated here to  $G_i$ , for the sake of brevity, in the substitutions that follow). Note

<sup>6</sup> Note that in the implementation of GALILEO there are two independent theories for the Dark Matter and the MOND case studies. For the sake of simplicity, the Isabelle theory for MOND was stripped down to the minimal structure needed to run and test the ORP rather than prove that the two theories derive from the same core.

$$S_{5,6} ::= S_{3,4} \cup \{Accel:Obj \Rightarrow Evt \Rightarrow \mathbb{R} \dots\} \quad (76)$$

$$S_5 ::= S_3 \quad (77)$$

$$A_5 ::= \{(48), (49), (50), (51), (52)\} \cup \{\forall e:Evt, s:Str, g:Gly. Accel(s, e) = \frac{(OV(s, g, e))^2}{Rad(s, g)}\} \quad (78)$$

$$S_6 ::= S_4 \quad (79)$$

$$A_6 ::= \{(54), (55), (56)\} \cup \{\forall e:Evt, s:Str, g:Gly. Accel(s, e) = \frac{(RtV(s, g, e))^2}{Rad(s, g)}, \quad (80)$$

$$c = 299792458\} \quad (81)$$

$$A_{6,1} ::= \{RedShift(Star1, Glxy71, Obs1) = 300 \quad (82)$$

$$Rad(Star1, Glxy71) = 2\} \quad (83)$$

⋮

$$A_{6,9} ::= \{RedShift(Star9, Glxy71, Obs9) = 300 \quad (84)$$

$$Rad(Star9, Glxy71) = 18\} \quad (85)$$

$$Sig(B_{O_{6,1}', O_5'}) ::= S'_{6,1} \cup S_5' \quad (86)$$

$$Ax(B_{O_{6,1}', O_5'}) ::= \{(61)\} \quad (87)$$

$$Sig(B_{O_{6,9}', O_5'}) ::= S'_{6,9} \cup S_5' \quad (88)$$

$$Ax(B_{O_{6,9}', O_5'}) ::= \{(61)\} \quad (89)$$

Fig. 7: *Initial model of the MOND case study*. Ontologies  $O_5'$  is a sub-ontology of  $O_5$  with only axioms (48), (49), (50), and (78). Bridges  $B_{O_{6,1}', O_5'}$  and  $B_{O_{6,9}', O_5'}$ , like  $B_{O_3', O_4'}$  in Figure 5, equate orbital and rotational velocity and allow to calculate the local value for  $G$ , which varies as the variad, i.e., the centripetal acceleration of a star (i.e., the rate of change of a star's tangential velocity around the galaxy). This allows the contradiction with  $O_5$  to become explicit.

also that  $O_5''$  is the sub-ontology in the factorized representation of  $O_5$  that contains all axioms of  $O_5$ :

$$O_5'' \vdash G ::= 6.673 \times 10^{-11} \quad (90)$$

$$O_{6,1}'(Accel(Star1, Obs1) = A_1) \oplus B_{O_{6,1}', O_5'} \oplus O_6' \oplus O_5' \vdash G = \frac{(OV(Star1, Glxy71, Obs1))^2 \times Rad(Star1, Glxy71)}{\sum s \in Glxy71/Star1.Mass(s, Obs1)} = G_1 \quad (91)$$

⋮ ⋮



$$\begin{aligned}
& \vdots \quad \quad \quad \vdots \\
O_{6,9}'(Accel(Star9, Obs9) = A_2) \oplus B_{O_{6,9}', O_5'} \\
& \oplus O_6' \oplus O_5' \vdash G = \\
& \frac{(OV(Star9, Glxy71, Obs9))^2 \times Rad(Star9, Glxy71)}{\sum_{s \in Glxy71/Star9.Mass(s, Obs9)}} \\
& = G_9 \tag{92}
\end{aligned}$$

$$O_5'' \vdash G_i \neq G_j \text{ for } 1 \leq i < j \leq 9 \tag{93}$$

Given the substitutions:

$$\begin{aligned}
\forall i \in [1, n] \\
\{O_5''/O_x, O_{6,i}'/O_{y_i}, B_{O_{6,i}', O_5'} \oplus O_6' \oplus O_5'/O_{z,i}, \\
G/stuff, \lambda x.x/f, 6.673 \times 10^{-11}/c, G_i/c_i, \\
Accel/v, \langle Star_i, Observation_i \rangle / \mathbf{b}_i, A_i/v_i\}
\end{aligned}$$

where  $O_6'$  is a sub-ontology in the factorised representation of  $O_6$  containing all of the original axioms. *Inconstancy* repairs the ten ontologies by redefining the inconstancy  $G$  as a function of  $G$  and of the variad  $Accel$  and by replacing the old definition with the new one:

$$\begin{aligned}
Ax(v(O_5'')) ::= \{\phi\{G/v(G)(\langle s \rangle)\} \mid \phi \in Ax(O_5)\} \\
\setminus \{G ::= 6.673 \times 10^{-11}\} \cup \\
\{v(G) ::= \lambda \langle s \rangle. F(G, Accel(\langle s \rangle))\} \tag{94}
\end{aligned}$$

$$\begin{aligned}
Ax(v(O_{6,i}'(Accel(\langle Star_i, Observation_i \rangle) = A_i))) ::= \\
\{\phi\{G/v(G)(\langle Star_i \rangle)\} \mid \phi \in \\
Ax(O_{6,i}'(Accel(\langle Star_i, Observation_i \rangle) = A_i))\} \tag{95}
\end{aligned}$$

$$\begin{aligned}
Ax(v(B_{O_{6,i}', O_5'} \oplus O_6')) ::= \{\phi\{G/v(G)(\langle s \rangle)\} \mid \\
\phi \in Ax(B_{O_{6,i}', O_5'} \oplus O_6')\} \tag{96}
\end{aligned}$$

The effect of the repair is not limited to the sub-ontologies indicated above, but it can be propagated to the entire corresponding representation. Every sub-ontology in the factorised representation receives the same treatment of repair.

## 5.2 Inconstancy application to the speed of light

One of the earliest recorded discussions of the speed of light,  $c$ , was by Aristotle, who believed that light travelled instantaneously and rejected theories about finite speeds of light. In 1676, a Danish astronomer, Ole Roemer, measured  $c$  by studying Io, one of Jupiter's moons, which was known to be eclipsed by Jupiter at regular intervals (Ellis and Uzan, 2005). Roemer discovered that the eclipses kept lagging behind the predicted times, but then started to pick up again. This discovery helped him come up with the theory that when Jupiter and Earth were further apart, there was more distance for light reflecting off Io to travel to Earth and therefore it took longer to reach his telescope. Figure 8 represents

$$\begin{aligned}
S_{13,14} ::= \{ \\
Light:Obj, Jupiter:Planet \dots \} \tag{97}
\end{aligned}$$

$$\begin{aligned}
S_{13} ::= \{ \\
TravTime:Obj \Rightarrow Time, \tag{98} \\
TravVel:Obj \Rightarrow Obj \Rightarrow Obj \Rightarrow \mathbb{R}, \dots \} \tag{99}
\end{aligned}$$

$$\begin{aligned}
A_{13} ::= \{ \\
\forall o_1, o_2, o_3:Obj, t:Time. TravVel(o_1) = \\
\frac{Dst(Pos(o_2, t), Pos(o_3, t))}{TravTime(o_1)}, \tag{100} \\
TravTime(Light) = 0 \} \tag{101}
\end{aligned}$$

$$\begin{aligned}
S_{14,1} ::= \{ \\
ArrTime:Obj \Rightarrow Time, \tag{102} \\
ObsVel:Obj \Rightarrow Obj \Rightarrow Obj \Rightarrow \mathbb{R}, \dots \} \tag{103}
\end{aligned}$$

$$A_{14,1} ::= \{ \tag{104}$$

$$Pos(Jupiter, 1) \neq Pos(Jupiter, 5) \tag{105}$$

$$Pos(Earth, 1) = Pos(Earth, 5) \tag{106}$$

$$\begin{aligned}
\forall o_1, o_2, o_3:Obj, t:Time. ObsVel(o_1, o_2, o_3, t) = \\
\frac{Dst(Pos(o_2, t), Pos(o_3, t))}{ArrTime(o_1) - t}, \tag{107}
\end{aligned}$$

$$Dst(Pos(Earth, 1), Pos(Jupiter, 1)) = 1, \tag{108}$$

$$ArrTime(Light) = 2 \} \tag{109}$$

$$S_{14,2} ::= \{(102), (103)\} \tag{110}$$

$$\begin{aligned}
A_{14,2} ::= \{ \\
(105), (106), (107), \tag{111} \\
Dst(Pos(Earth, 5), Pos(Jupiter, 5)) = 5, \tag{112} \\
ArrTime(Light) = 7 \} \tag{113}
\end{aligned}$$

$$Sig(B_{O_{13}', O_{14,1}'}) ::= S_{13} \cup S_{14,1} \tag{114}$$

$$Ax(B_{O_{13}', O_{14,1}'}) ::= \{TravVel = ObsVel\} \tag{115}$$

$$Sig(B_{O_{13}', O_{14,2}'}) ::= S_{13} \cup S_{14,2} \tag{116}$$

$$Ax(B_{O_{13}', O_{14,2}'}) ::= \{TravVel = ObsVel\} \tag{117}$$

Fig. 8: *Initial model of the Light case study.* Ontologies  $O_{13}'$ ,  $O_{14,1}'$  and  $O_{14,2}'$  are sub-ontologies of  $O_{13}$ ,  $O_{14,1}$ , and  $O_{14,2}$ , respectively.  $O_{13}'$  contains only the axiom (100),  $O_{14,1}'$  contains all the axioms of  $O_{14,1}$ , and  $O_{14,2}'$  contains all the axioms of  $O_{14,2}$ . Bridges  $B_{O_{13}', O_{14,1}'}$  and  $B_{O_{13}', O_{14,2}'}$  allow to calculate the local value for travel time  $TravTime$ , which varies as the variad. This allows the contradiction with  $O_{13}$  to become explicit.

the state of the theory of the speed of light at the time of Rømer's measurement. At the top  $S_{13,14}$  contains the shared signature with type definitions, for instance, for *Light* and *Jupiter*.  $O_{13}$  defines in equation (100) the velocity *TravVel* of an object  $o_1$  from  $o_2$ 's position to  $o_3$ 's position as the ratio between the distance between  $o_2$  and  $o_3$  and  $o_1$ 's travel time, *TravTime*.  $O_{13}$  also contains Aristotle's postulate (101) according to which light's *TravTime* is zero (making the result of equation (100) impossible). On the other hand,  $O_{14.1}$  and  $O_{14.2}$  assume that *Jupiter* has different positions in times 1 and 5 and in equation (107) it defines the observed velocity *ObsVel* of an object  $o_1$  travelling from  $o_2$ 's to  $o_3$ 's position as the ratio between the distance from  $o_2$ 's to  $o_3$ 's position at the particular time of  $o_1$ 's departure and the difference between  $o_1$ 's time of arrival and its time of departure.  $O_{14.1}$  and  $O_{14.2}$  contain estimates for the distances between *Earth* and *Jupiter* at times 1 respectively 5, as well as measurements for light's *ArrTime*. Finally, the bridges <sup>7</sup>  $B_{O_{13},O_{14.1}}$  and  $B_{O_{13},O_{14.2}}$  link together the seemingly disparate terms by relating *TravVel* and *ObsVel*. This allows the contradiction about the value of *TravTime(Light)* between  $O_{14.1}$  and  $O_{14.2}$  to become explicit, as follows:

$$O_{13}'' \vdash \text{TravTime(Light)} = 0 \quad (118)$$

$$O_{14.1}' (\text{Dst(Pos(Earth, 1), Pos(Jupiter, 1)) = 1}) \oplus B_{O_{13}',O_{14.1}'} \oplus O_{13}' \vdash \quad (119)$$

$$O_{14.2}' (\text{Dst(Pos(Earth, 5), Pos(Jupiter, 5)) = 5}) \oplus B_{O_{13}',O_{14.2}'} \oplus O_{13}' \vdash \quad (120)$$

$$O_{13}'' \vdash 1 \neq 2 \quad (121)$$

where  $O_{13}''$  is a sub-ontology in the factorised representation of  $O_{13}$  containing all of the original axioms such that the above holds. Given the substitutions

$$\begin{aligned} & \{O_{13}''/O_x, O_{14.i}'/O_{y_i}, B_{O_{13}',O_{14.i}'} \oplus O_{13}'/O_{z_i}, \\ & \text{stuff}/\text{TravTime(Light)}, \lambda x.x/f, c/0, \\ & c_i/\text{TravTime(Light)}_i, v/\text{Dst}, \\ & \mathbf{b}_i/\langle \text{EarthPos}, \text{Pos(Jupiter}, t_i) \rangle, v_i/D_i \} \end{aligned}$$

*Inconstancy* repairs the three ontologies by redefining the inconstancy *TravTime(Light)* as a function of *TravTime(Light)* and of the variad *Dst* and by replacing the old definition with

the new one:

$$\begin{aligned} Ax(v(O_{13}'')) ::= & \\ & \{ \phi \{ \text{TravTime(Light)} / \\ & \quad v(\text{TravTime(Light))(\langle p_m, p_n \rangle) \} \mid \\ & \phi \in Ax(O_{13}'') \} \setminus \\ & \{ \text{TravTime(Light)} = 0 \} \cup \\ & \{ v(\text{TravTime(Light)}) ::= \lambda \langle p_m, p_n \rangle. \\ & \quad F(\text{TravTime(Light)}, \text{Dst}(\langle p_m, p_n \rangle)) \} \quad (122) \end{aligned}$$

$$\begin{aligned} Ax(v(O_{14.1}')) & \\ & (\text{Dst}(\langle \text{Pos(Earth, 1)}, \text{Pos(Jupiter, 1)} \rangle) = 1)) ::= \\ & \{ \phi \{ \text{TravTime(Light)} / v(\text{TravTime(Light)}) \\ & \quad (\langle \text{Pos(Earth, 1)}, \text{Pos(Jupiter, 1)} \rangle) \} \mid \phi \in \\ & \quad Ax(O_{14.1}' (\text{Dst}(\langle \text{Pos(Earth, 1)}, \text{Pos(Jupiter, 1)} \rangle) \\ & \quad = 1)) \} \quad (123) \end{aligned}$$

$$\begin{aligned} Ax(v(O_{14.2}')) & \\ & (\text{Dst}(\langle \text{Pos(Earth, 5)}, \text{Pos(Jupiter, 5)} \rangle) = 5)) ::= \\ & \{ \phi \{ \text{TravTime(Light)} / v(\text{TravTime(Light)}) \\ & \quad (\langle \text{Pos(Earth, 5)}, \text{Pos(Jupiter, 5)} \rangle) \} \mid \phi \in \\ & \quad Ax(O_{14.2}' (\text{Dst}(\langle \text{Pos(Earth, 5)}, \text{Pos(Jupiter, 5)} \rangle) \\ & \quad = 5)) \} \quad (124) \end{aligned}$$

$$\begin{aligned} Ax(v(B_{O_{13}',O_{14.i}'} \oplus O_{13}')) ::= & \\ & \{ \phi \{ \text{TravTime(Light)} / \\ & \quad v(\text{TravTime(Light))(\langle p_m, p_n \rangle) \} \mid \\ & \phi \in Ax(B_{O_{13}',O_{14.i}'} \oplus O_{13}') \} \quad (125) \end{aligned}$$

The effect of the repair is not limited to the sub-ontologies indicated above, but it can be propagated to the entire corresponding representation. Every sub-ontology in the factorised representation receives the same treatment of repair.

## 6 The *Unite* ORP

The *Unite* ORP, described in Figure 9, is triggered when the observed values of the defining properties of two distinct  $f(\text{stuff})$ s are the same<sup>8</sup>. Just like *WMS*, *Unite* assumes two ontologies: an ontology  $O_x$  representing a predictive Physics theory and labelling a given property as defining and another ontology  $O_y$  representing some sensory information arising from observations. In case of a heterogeneous representation, a bridging ontology may be involved in the reasoning. Contrary to *WMS*, *Unite* is not triggered by a contradiction, but by an equality of two  $f(\text{stuff})$ s (a redundancy) relative

<sup>7</sup> Note that in the implementation of *GALILEO* there are no bridges for the *Light* case study. For the sake of simplicity, the *Isabelle* theory was stripped down to the minimal structure (i.e., the same signature was used for both ontologies) to test the ORP.

<sup>8</sup> For instance, the orbits of (heavenly) bodies is one of their defining properties, i.e., equating two orbits is equivalent to identifying the corresponding bodies as the same one, according to the principle that two objects cannot be at the same place at the same time.

Trigger:  $O_x$  is the predictive ontology and  $O_y$  is the observational one.

If  $f(stuff_1)$  and  $f(stuff_2)$  cannot be proven to have the same nor different values in  $O_x$  and in  $O_y$ , but the value of  $f(stuff_1)$  and  $f(stuff_2)$  for meta-level property  $dp$ , labelled as defining in the meta-level ontology  $O_M$ , is the same, then the following formula will be triggered:

$$\exists O_x, O_y: \mathbb{O}, \tau, \tau', \tau'': Types, dp: \tau \Rightarrow \tau', stuff_1, stuff_2: \tau$$

$$f: \tau \Rightarrow \tau'$$

$$(O_x \not\vdash f(stuff_1) =_{\tau'} f(stuff_2)) \quad (126)$$

$$O_x \not\vdash f(stuff_1) \neq_{\tau'} f(stuff_2) \quad (127)$$

$$O_M \vdash DefProp(dp, \tau') \quad (128)$$

$$O_y \vdash dp(stuff_1) =_{\tau'} dp(stuff_2) \quad (129)$$

where  $DefProp(dp, \tau')$  means that the function  $dp$  is a defining property for objects of type  $\tau'$ .

Repair: Since the two  $f(stuff)$ s have the same value for its defining property of objects of their type, they are identified as one.

$$f(stuff_1) =_{\tau} f(stuff_2) \quad (130)$$

Let  $v(O_x)$  be the repaired predictive ontology. The axioms for the new ontologies are updated in terms of those of the old, where  $Ax(O)$  returns all axioms of  $O$ , as follows:

$$Ax(v(O_x)) ::= \{f(stuff_1) =_{\tau} f(stuff_2)\} \cup Ax(O_x) \quad (131)$$

To effect the repair, the axioms of  $v(O_x)$  are the same as those of  $O_x$  except for the addition of the new definition.

Fig. 9: The Unite ORP

to the chosen defining property. The repair is to equate in  $O_x$   $stuff_1$  and  $stuff_2$ .

### 6.1 Unite application to the Morning and Evening Star

Because Venus is closer to the Sun than the Earth, it becomes visible either just before dawn or just after sunset, when it is the brightest heavenly object after the Moon. These two kinds of appearance were not originally identified as coming from the same object. It was only with the quantification of astronomy that the 4-dimensional orbits<sup>9</sup> of these two stars were calculated and seen to be the same (up to experimental error).

One way of comparing the two orbits is to calculate their *mean anomaly*  $M$ , i.e., the parameter relating position and time for a body moving in a Kepler orbit:

$$M = \sqrt{\frac{G(M+m)}{a^3}} \times t \quad (132)$$

where  $a$  is the length of the orbit's semi-major axis,  $M$  and  $m$  are the orbiting masses, and  $G$  is the gravitational constant,  $t$  is time.

For instance the mean anomaly of specific bodies orbiting the Sun can be understood as the time since the last point

<sup>9</sup> The 4-dimensional orbits covers the three dimensions of space and time

of closest approach to the Sun (periapsis) multiplied by their mean motion.

Figure 10 axiomatizes the case study at hand.  $S_{9,10}$  provides type declarations for the Morning Star,  $MS$ , the Evening Star,  $ES$ , and the Sun,  $Sun$ , which are of the type of objects  $Obj$ .  $O_9$  provides theoretical knowledge for calculating the mean anomaly of objects orbiting the Sun and declares the defining property.  $O_9$  asserts only definitions but no data, because no prediction is made using the new heliocentric theory in its own right. On the other hand,  $O_{10}$  contains no definitions, as the old geocentric theory is not relevant, while the observations of  $MS$  and  $ES$  in  $O_{10}$  are. The observations of  $MS$  and  $ES$  are based on the angles of the visual range and the distances from Earth. The bridging ontology<sup>10</sup>  $B_{O_9, O_{10}}$  relates together the observed angles of an object, the distance of the object from Earth and the radius of the object.

$$O_9' \not\vdash MS = ES \quad (151)$$

$$O_9' \not\vdash MS \neq ES \quad (152)$$

$$O_M \vdash DefProp(MeanAnom, Obj) \quad (153)$$

$$O_9' \oplus B_{O_9', O_{10}'} \oplus O_{10}' \vdash \forall e. MeanAnom(MS, e) = MeanAnom(ES, e) \quad (154)$$

Given the substitutions:

$$\{O_9'/O_x, O_9' \oplus B_{O_9', O_{10}'} \oplus O_{10}'/O_y, MS/stuff_1, ES/stuff_2, MeanAnom/dp, \lambda x.x/f\} \quad (155)$$

Unite repairs  $O_9'$  as follows:

$$Ax(v(O_9')) ::= \{MS =_{\tau} ES\} \cup Ax(O_9')$$

The effect of the repair is not limited to  $v(O_9')$  but applies to the entire corresponding representation. Every sub-ontology in the representation receives the same treatment of repair.

### 6.2 Unite application to the shape of the Earth

Pythagoras was one of the first astronomers to realise that the Earth was a sphere. He gathered evidence to support this theory from various sources, among which were observations of lunar eclipses. He noticed that the edge of the shadow that the Earth cast on the Moon was always circular. He reasoned that the only 3D shape that always casts circular shadows is a sphere, which is also the shape of a ball.

The defining property for shapes of objects is

$$\lambda x. ShapeofParam(Project(x, Flat)) \quad (165)$$

which is essentially the shape of the projection of an object  $x$  onto a flat surface. The idea is that if two 3D objects always

<sup>10</sup> Note that in the implementation of GALILEO there are no bridges for the Morning and Evening Star case study. For the sake of simplicity, the Isabelle theory was stripped down to the minimal structure (i.e., the same signature was used for both ontologies) to test the ORP.

$$S_{9,10} ::= \{MS:Obj, ES:Obj, Sun:Obj \dots\} \quad (133)$$

$$S_9 ::= \{MeanAnom:Obj \Rightarrow Evt \Rightarrow \mathbb{R}Set, \dots\} \quad (134)$$

$$A_9 ::= \{\forall o:Obj, e:Evt. MeanAnom(o, e) = \sqrt{\frac{G \times (Mass(Sun, e) + Mass(o, e))}{Rad_a^3(o, Sun, e)}} \times Duration(e)\} \quad (135)$$

$$A_M ::= \{DefProp(MeanAnom, Obj)\} \quad (136)$$

$$S_{10} ::= \{\} \quad (137)$$

$$A_{10} ::= \{\forall e:Evt. Mass(Sun, e) = 1.9891 \times 10^{30}, \quad (138)$$

$$\forall e:Evt. Mass(MS, e) = 4.868 \times 5 \times 10^{24}, \quad (139)$$

$$\forall e:Evt. Mass(ES, e) = 4.868 \times 5 \times 10^{24}, \quad (140)$$

$$Angle(MS, Obs1) = 1.3 \times 10^{-4}, \quad (141)$$

$$Angle(ES, Obs1) = 1.3 \times 10^{-4}, \quad (142)$$

$$Angle(MS, Obs5) = 1.2 \times 10^{-4}, \quad (143)$$

$$Angle(ES, Obs5) = 1.2 \times 10^{-4} \quad (144)$$

$$Dst(MS, Earth, Obs1) = 44.7 \times 10^6, \quad (145)$$

$$Dst(ES, Earth, Obs1) = 44.7 \times 10^6, \quad (146)$$

$$Dst(MS, Earth, Obs5) = 45.5 \times 10^6, \quad (147)$$

$$Dst(ES, Earth, Obs5) = 45.5 \times 10^6 \quad (148)$$

$$Sig(B_{O_9', O_{10}'} ) ::= S_9 \cup S_{10} \cup S_{9,10} \quad (149)$$

$$Ax(B_{O_9', O_{10}'} ) ::= \{\forall o:Obj, e:Evt, v:\mathbb{R}. Rad(o, e) = Dst(o, Earth, e) \times \tan(\frac{1}{2}Angle(o, e))\} \quad (150)$$

Fig. 10: *Initial model of the Morning and Evening Star case study.* Ontologies  $O_9'$  and  $O_{10}'$  are two sub-ontologies of  $O_9$  and  $O_{10}$ , respectively.  $O_9'$  contains all the axioms of  $O_9$  and  $O_{10}'$  contains all the axioms of  $O_{10}$ . The bridge  $B_{O_9', O_{10}'}$  allows to calculate the radius of  $MS$  and  $ES$  from measurements of the angle of visual range and their distances from Earth. With  $B_{O_9', O_{10}'}$ , the definitions and data available in  $O_9$  and  $O_{10}$  can be used to identify that the two stars are one heavenly body.

$$S_{11,12} ::= \{Shapeof:Obj \Rightarrow Shape, Earth:Obj, Ball:Obj, Flat:Surface \dots\} \quad (156)$$

$$S_{11} ::= \{Project:Obj \Rightarrow Surface \Rightarrow (\mathbb{R} \Rightarrow \mathbb{R} \times \mathbb{R}), \quad (157)$$

$$ShapeofParam:(\mathbb{R} \Rightarrow \mathbb{R} \times \mathbb{R}) \Rightarrow Shape, \dots\}$$

$$A_{11} ::= \{ShapeofParam(a, b) = circle \leftrightarrow \exists k.a = (\lambda m.k \times \cos(m)) \wedge b = (\lambda m.k \times \sin(m))\} \quad (158)$$

$$A_M ::= \{DefProp(\lambda x.ShapeofParam(Project(x, Flat)), Shape)\} \quad (159)$$

$$S_{12} ::= \{\} \quad (160)$$

$$A_{12} ::= \{Project(Earth, Flat) = \lambda t.(2 \times \cos(t), 2 \times \sin(t)), \quad (161)$$

$$Project(Earth, Flat) = \lambda t.(2 * 10^{10} * \cos(t), 2 * 10^{10} * \sin(t))\} \quad (162)$$

$$Sig(B_{O_{11}', O_{12}'}) ::= S_{11} \cup S_{12} \cup S_{11,12} \quad (163)$$

$$Ax(B_{O_{11}', O_{12}'}) ::= \{\} \quad (164)$$

Fig. 11: *Initial model of the shape of the Earth case study.* Ontologies  $O_{11}'$  and  $O_{12}'$  are two sub-ontologies of  $O_{11}$  and  $O_{12}$ , respectively.  $O_{11}'$  contains all the axioms of  $O_{11}$  and  $O_{12}'$  contains all the axioms of  $O_{12}$ .  $O_{11}' \oplus B_{O_{11}', O_{12}'} \oplus O_{12}'$  allows to calculate the shape of *Earth* and *Ball* from the definitions available in  $O_{11}'$  and the data available in  $O_{12}'$ ; thus, we can identify the shapes of the two projections being the same.

have the same 2D projections then they have the same shape themselves. For the projection of the Earth onto a (almost) flat surface, e.g., the surface of the Moon, the parameterisation of the shape of the projection is

$$\lambda t.(2 \times 10^{10} \times \cos(t), 2 \times 10^{10} \times \sin(t)) \quad (166)$$

and for a ball onto a flat surface, the parameterisation is

$$\lambda t.(2 \times \cos(t), 2 \times \sin(t)) \quad (167)$$

Thus, the parameterisations of the projections differ only in scale. Multiple, independent projections are required in order to draw an accurate conclusion. A cylinder also projects as a circle along its axis, but most of its projections are not circular, so one projection is not enough. Note also that the observed projections of *Ball* are a thought experiment.

Figure 11 axiomatizes the case study at hand.  $S_{11,12}$  provides type declarations for, among others, the symbols *Shapeof* and *Earth*.  $A_{11}$  provides part of the theoretical knowledge

for orthographic projections.  $A_M$  declares the defining property for such knowledge. On the other hand,  $A_{12}$  contains observations of lunar eclipses. Given  $O_{11}'$  and  $O_{12}'$  each being a sub-ontology of  $A_{11}$  and of  $A_{12}$ , respectively,  $O_{11}' \oplus B_{O_{11}', O_{12}'} \oplus O_{12}'$  allows to apply the theory to the available data as follows:

$$O_{11}' \not\vdash \text{Shapeof}(\text{Earth}) = \text{Shapeof}(\text{Ball}) \quad (168)$$

$$O_{11}' \not\vdash \text{Shapeof}(\text{Earth}) \neq \text{Shapeof}(\text{Ball}) \quad (169)$$

$$O_M \vdash \text{DefProp}(\lambda x. \text{ShapeofParam}(\text{Project}(x, \text{Flat})), \text{Shape}), \quad (170)$$

$$O_{11}' \oplus B_{O_{11}', O_{12}'} \oplus O_{12}' \vdash \\ \text{ShapeofParam}(\text{Project}(\text{Earth}, \text{Flat})) = \\ \text{ShapeofParam}(\text{Project}(\text{Ball}, \text{Flat}))$$

Given the substitutions:

$$\{O_{11}'/O_x, O_{11}' \oplus B_{O_{11}', O_{12}'} \oplus O_{12}'/O_y, \text{Earth}/\text{stuff}_1, \\ \text{Ball}/\text{stuff}_2, \lambda x. \text{ShapeofParam}(\text{Project}(x, \text{Flat}))/dp, \\ \text{Shapeof}/f\}$$

*Unite* repairs  $O_{11}'$  as follows:

$$Ax(v(O_{11}')) ::= \{\text{Shapeof}(\text{Earth}) = \text{Shapeof}(\text{Ball})\} \cup \\ Ax(O_{11}')$$

The effect of the repair is not limited to  $v(O_{11}')$  but to the entire corresponding representation. Every sub-ontology in the representation receives the same treatment of repair.

## 7 Implementation

GALILEO is prototyped in the higher-order theorem prover Isabelle (Paulson, 1994). Ontologies and bridges are represented as *Locales* (Ballarin, 2004), i.e., as independent proof contexts that can be extended. If a locale  $M$  is an extension of a locale  $L$ , then all symbols in the signature of  $L$  are in the signature of  $M$  and all theorems of  $L$  are theorems of  $M$ , as expected. With the ability to create extensions, merging locales is naturally possible, as merging two locales can be obtained by creating a new locale that extends from both of them. For our application, extensions essentially allow to modularise and express dependencies between the ontologies.

Just like the corresponding ORPs illustrated in Fig. 3, 6, 9, each of GALILEO's procedures supports two main reasoning tasks: the diagnosis of the conflict between the considered ontologies and their repair.

### 7.1 Conflict diagnosis

The goal of conflict diagnosis is the correct instantiation of the term *stuff* as well as of every existential variable in the

trigger. This is essentially a reasoning task, which is performed by using Isabelle's higher-order matching (HOM) algorithm, a special case of higher-order unification (HOU). In each ORP, conflict diagnosis consists of four main phases: preparation, verification, term discovery, heuristics-based filtering.

#### 7.1.1 Preparation

In order to reduce the inferential search scope, the user provides the conflicting ontologies (e.g.,  $O_1'$  and  $B \oplus O_2'$ ) as well as the *shape* of the contradiction, i.e., a sentence that is provable in one of the ontologies but not in the other (e.g.,  $\text{Heat}(\text{H}_2\text{O}, \text{Melting}) = 0$ ). Note that the ontologies indicated here are expected to be sub-ontologies of the input, i.e. ontologies within the factorised representation. This preparation phase is made necessary by the fact that the instantiation of existentially quantified higher order variables is not tractable in the highly general setting we are working with. Therefore the contradiction between two or more given ontologies must be made explicit. In practice, the higher order theorem prover is provided with a theory file specifying the ontologies as well as the proofs of the conflicting statements that constitute the trigger of each ORP.

Note that the limitation on the instantiation of existentially quantified higher order variables, imposes an additional information requirement in the case of *Unite*. As explained, to avoid looping, this ORP requires the instantiation of  $\text{stuff}_1$  and  $\text{stuff}_2$  rather than the instantiation of  $f(\text{stuff}_1)$  and  $f(\text{stuff}_2)$ . Therefore during the preparation phase *Unite* is provided with the instantiations of  $\text{stuff}_1$  and  $\text{stuff}_2$ .

#### 7.1.2 Verification

Before proceeding to the analysis of the terms that appear in the contradictory statements, a verification procedure checks that the contradictory statements indeed follow from the provided ontologies. Technically this is done by checking that the contradictory statements can be discharged as proof obligations in the search space specified in the theory file. This phase is important for ensuring that the appropriate repair is to be performed on the faulty ontologies.

#### 7.1.3 Term discovery by higher-order matching

Since ORPs are intended to be highly general, GALILEO automatically determines the correct concrete type of the existential variables in the corresponding trigger formula, including  $f$ ,  $\text{stuff}$ , and  $v$ . For instance, in the Latent Heat case study presented in Fig. 4,  $\text{Heat}(\text{H}_2\text{O}, \text{Melting}) > 0$  is a sentence that is provable in one ontology (e.g.  $O_1$ ) but not in another ( $O_2$ ). Here the exact instantiation of the variable  $\text{stuff}$  is not immediately clear because  $\text{stuff}$  is polymorphic and

variadic. The sentence  $Heat(H_2O, Melting) > 0$  is an object-level formula, so the identification of the instantiations of *stuff* must be a meta-level procedure. Using schematic-variables (or meta-variables)  $?f$ ,  $?stuff$ ,  $?v$ , and  $?op$ , the following pattern can be created:

$$?op(?f(?stuff), ?v) \quad (171)$$

where  $?f$ ,  $?stuff$ , and  $?v$  respectively corresponds to  $f$ ,  $stuff$ , and  $v$  in WMS's presentation in Figure 3, and  $?op$  corresponds to the function used for comparing between  $f(stuff)$  and  $v$ <sup>11</sup>. So, to obtain the instantiation for each existential variable in the trigger formulae, the complete pattern for, e.g., WMS, is in the following shape:

$$\exists v_1, v_2. ?O \implies ?op_1(?f(?stuff), v_1) \wedge ?op_2(v_1, v_2) \dots \quad (172)$$

where  $?O$  is expected to be instantiated to an ontology in a relevant factorised representation, and  $op_1$  and  $op_2$  correspond to partial orders. Here, ontologies are assumed to be predicates, where each axiom is a conjunct. Patterns based on (171) actually hit a highly difficult aspect of HOU, as matching it against a sentence demands the creation of possible functions to solve a constraint involving a function variable. Isabelle imposes a restriction on the automation such that a meta-type variable  $?α$  is not transformed into one of some function type. Without this restriction, each free variable can be unified to many more terms, yielding a unification space too large for practical applications. However, since  $?stuff$  is supposed to be variadic and polymorphic, assigning the type of  $?stuff$  to  $?α$  is too restrictive for our application. For instance, if we match (171) with  $?stuff : ?α$  against the sentence  $Heat(H_2O, Melting) > 0$  using HOM,  $stuff \Rightarrow Heat$  would not be a valid substitution, because  $Heat$  is a second-order function with an arity of two. Our solution is to generate a collection of type patterns that range over the potential arities and function order of  $stuff$  and  $v$ . Each type pattern is then used for matching (171) against the sentence, and therefore,  $stuff$  can be instantiated to functions of different types with different arities. Similarly,  $?v$  can also be instantiated to functions of various arities. For instance, the Latent Heat case study requires  $?v$  to be instantiated to 0, which is of type  $\mathbb{R}$ , whereas the Dark Matter case study requires it to be instantiated to  $UpFlat$ , which is a function of type  $\mathbb{R} \Rightarrow \mathbb{R}$ .

Furthermore, as shown in Tables 2, 3, 4, the depth of the unification is determined by a so-called unify search bound (*usb*), which effectively acts as a limit on the logical order of the matches returned by the system. For instance, in Table 2 for  $3 \leq usb \leq 4$  the system returns zero<sup>th</sup> order matches (i.e. propositional matches), for  $usb = 5$  the system returns first order matches, for  $usb = 6$  the system returns second order matches, etc. Note also that the distribution of matches

across the various *usb*'s from top to bottom provides insight into the system's reasoning trace for what concerns attempted generalizations. However, the relation between *usb* and logical order is not absolute, it varies per case study, as it depends on the term in question. Also note that, for the same case study, the *usb*'s at which a given set of solutions is found may vary with the version of GALILEO. That is because the *usb* also depends on the pattern being matched and new patterns are added to new versions.

#### 7.1.4 Heuristics-based filtering

Due to the high generality of the trigger formulae determined by  $f$ 's and  $stuff$ 's polymorphism, the space of possible instantiations typically contains a large number of matches. These, though, often contain arbitrary  $\lambda$ -expressions that have no meaningful physical relevance. A heuristics-based filtering is used to automatically prune the solution space, e.g., disregarding instantiations that contain either:

- the identity function;
- no element of the signature;
- the same functional symbol applied to permuted arguments;
- a  $\lambda$ -abstracted variable as dominant function.

## 7.2 Repair

Once a match is selected, the repair takes place. Recall that the existentials representing the ontologies in conflict are instantiated to some ontologies in a factorised representation of the originals in order to trigger an ORP. However, the effect of repair is not limited to the instantiations but to the entire corresponding representation. Every ontology in a representation receives the same treatment of repair if deemed relevant, e.g., if it contains occurrences of the instantiation of *stuff*.

Occurrences of the instantiation of *stuff* in an axiom are checked by HOM as well. Even though the transformation rules of repair require only syntactic analysis of the axioms, formalising repair as a HOM problem provides more robustness to the whole procedure. Thus, the look-up of occurrences of *stuff* in a formula involves searching for a match.

## 8 Evaluation

As anticipated in Section 1, the empirical results provided by the evolutions proposed by GALILEO allow to confirm our hypothesis regarding the evaluation of various desirable properties: coverage, efficiency, maintainability, high quality of the repairs. In particular, the results allow to conclude that the generality of a Higher-Order Logical approach to

<sup>11</sup> Note that partial orders can be expressed as binary functions.

ontology evolution is beneficial to the quality of the repairs, in terms of the meaningfulness of the proposed evolutions.

### 8.1 Coverage and efficiency

The expressivity of a polymorphic Higher-Order Logic is valuable in achieving generality. The instantiations of *stuff*, as well as of the other variables in the ORPs’s trigger formulae, are diverse in types and arities, and allow to handle disparate case studies.

On the other hand, exactly because of HOL’s expressivity, the initial number of matches returned by the theorem prover before the heuristics are applied is extremely high. When diagnosed at higher *usb*s (between 6 to 10), some case studies can return millions of instantiations for *stuff* and the other meta-variables. The heuristics successfully reject the inappropriate matches and make the set of solutions viable and correct, as the historically accurate solution is always present.

### 8.2 Maintainability

There are two aspects of GALILEO’s implementation that pose maintainability issues. On the one hand, there is the maintenance of the ORPs, i.e., of the HOM-based diagnostic and repair machinery. Here the main problems are the required knowledge of Isabelle’s programming language, ML, and the interaction between the ORPs and Isabelle. In order to make this interaction as easy to maintain as possible GALILEO is based on a modular architecture, where ORPs are independent of one another.

On the other hand, there is the modeling effort required for the case studies. Such modeling has proven to be a very demanding and time-consuming task, which requires a set of multidisciplinary skills ranging from an understanding of the Physics involved in a given case study to the appreciation of the historical state of the field at the time of the case study, from an eye for the conceptual and logical analysis needed for formalizing the case studies, to the practical ability at implementing knowledge and proofs in a theorem prover such as Isabelle. As further explained in Section 9, unfortunately we have not found any viable pre-existing formalisation of the required physical theories. Yet such a formalization would make implementing the case studies much easier and, more importantly, would facilitate deeper inference to extend the range of ontology repairs.

### 8.3 Quality of the repairs

The multiplicity of matches returned by the ORPs can be evaluated from two main perspectives:

<i>usb</i>	Match of <i>stuff</i>	No.
3,4	$H_2O:Obj$	(0)
3,4	$Melting:Evt$	(1)
3,4	$Heat(H_2O, Melting):\mathbb{R}$	(2)
5	$\lambda a. Heat(a, Melting):Obj \Rightarrow \mathbb{R}$	(3)
5	$\lambda a. Heat(H_2O, a):Evt \Rightarrow \mathbb{R}$	(4)
6	$\lambda a b. Heat(a, b):Obj \Rightarrow Evt \Rightarrow \mathbb{R}$	(5)

Table 2: Output of diagnosis for the Latent Heat case study at progressively larger unify search bounds  $0 \leq usb \leq 6$ . Before *usb* = 3, matches do not make it through the heuristics. The historically preferred match is (5), which is reached at *usb* = 6.

<i>usb</i>	Match of <i>stuff</i>	No.
5	$\lambda a. Glxy71 a:Obj \Rightarrow bool$	(0)
5	$\lambda a. GphB (\lambda a. Glxy71 a) a:\mathbb{R} \Rightarrow \mathbb{R}$	(1)

Table 3: Output of diagnosis for the Dark Matter case study at progressively larger unify search bounds  $0 \leq usb \leq 5$ . Before *usb* = 5, matches do not make it through the heuristics. The historically preferred match is (0), which is reached at *usb* = 5.

<i>usb</i>	Match of <i>stuff</i>	No.
2	$G:\mathbb{R}$	(0)

Table 4: Output of diagnosis for the MOND case study, at progressively larger unify search bounds  $0 \leq usb \leq 2$ . Before *usb* = 2, matches do not make it through the heuristics. The preferred instantiation is (0), which is reached at *usb* = 2.

1. Is the accurate historical match returned?
2. Do other matches correspond to physically interesting ontology repairs?

In all cases provided to GALILEO, the historically preferred diagnoses are returned. These are (5) in Table 2, (0) in Table 3, (0) in Table 4, (0) in Table 5. For what concerns *Unite*, as explained, this ORP is provided with the instantiations of *stuff*<sub>1</sub> and *stuff*<sub>2</sub>, and this makes the search unnecessary; therefore no table of diagnoses is given for this ORP.

For the other ORPs, though, a number of alternate matches make it through the heuristics. These matches can be classified, e.g., as experimental errors or as physically plausible alternate matches.

usb	Match of <i>stuff</i>	No.
2,3	$TravTime(Light):\mathbb{R}$	(0)
2,3	$TravTime(Light):\mathbb{R}$	(1)
0,1	$Light:Obj$	(2)

Table 5: *Output of diagnosis for the Light case study, at progressively larger unify search bounds  $0 \leq usb \leq 3$ . Before  $usb = 0$ , matches do not make it through the heuristics. The preferred instantiation is (0), which is reached at  $usb = 2$ . Note that matches reached at  $usb = 0$  are printed after those reached at a higher  $usb$ ; this reflects how the automated theorem prover’s search strategy is modified by increasing the  $usb$ , which allows it to explore more general solutions first. Also the redundancy of matches (0) and (1) is only apparent, as the instantiation of the other meta-variables (e.g.  $v$  or  $b_i$ ) does vary in the two diagnoses.*

Experimental errors are matches of the lowest logical order, such as (0), (1), (2) in Table 2, (0) in Table 4 and (2) in Table 5. Matches of this type do not seem to suggest the need for new physical laws, but rather the need to repeat the measurements because of experimental errors. In the case of (0) in Table 4, the suggested revision has a general value, in the sense that what is suggested is the revision of the value of a physical constant such as  $G$ , i.e., of a fundamental component of a Physics theory, which, though, is subject to measurement.

Physically plausible alternate matches suggest viable repairs leading to alternative physical laws. For instance, (3) and (4) in Table 2 are two such cases. (4) suggests to repair the heat function specific to water. As discussed in Sec 4.1 this solution was actually investigated by Black, and led to the definition of the notion of specific heat (constant  $c$ ). (3), on the other hand, suggests a repair to the heat function specific to melting. This solution too is viable, and actually led to the study of calorimetry.

#### 8.4 Degree of automation

Given the potentially vast number of ontologies generated in a factorised representation, it is essential to automate the factorisation of the input ontologies. GALILEO incorporates the ability to take an ontology as input and generate the sub-ontologies.

As outlined in Section 7.1, the execution of conflict diagnosis itself consists of three components: *verification*, which is a procedure for checking if an ORP is indeed applicable to the given ontologies; *term discovery*, which identifies all valid instantiations of the variable *stuff*; and, *heuristics-based filtering*, which eliminates physically implausible hy-

potheses (potential repairs). In the implementation, only the verification requires user interaction, because solving non-trivial reasoning problems in Higher-Order Logic typically requires manual guidance. Fortunately, GALILEO is implemented on top of Isabelle, so the user has access to all theorem proving facilities available in Isabelle, including automated first-order provers, a range of tactics, and a rich library of theorems. Once verification is completed, the system searches for all physically plausible potential repairs and returns them to the user. Typically, more than one potential repair remains after filtering and the user is required to select the most appropriate repair with respect to their specific intention. User interaction is required here as the system has no knowledge about the user’s intention or desire.

The repair process is entirely automated. The system analyses the parameters of both the dependencies and the morphisms between each ontology in the environment to decide in which ontology the repair should be effected. In all, the meta-level reasoning and operations are entirely automated, whereas some interactive guidance is required at the object-level due to the intrinsic complexity underlying HOL reasoning.

#### 8.5 Comparison with DL approaches

Another possible element of evaluation of the approach is a comparison with DL approaches. Compared to the approaches found in the literature and discussed in Section 2, our approach expands the space of solutions. The application of an Ontology Repair Plan often combines the retraction and/or addition of axioms as well as the deeper modification of the language in which the ontology is represented. Note that a version of WMS specialized for DL would have to produce repairs for subsumption axioms based on  $\sqsubseteq$  and on  $\sqcup$ , rather than for equations based on  $::=$  and on  $+$ , like equation (17). If such a specialized version of WMS were applied to the incoherent and inconsistent ontology network of university and company members shown in Figure 2, an additional class of solutions would be proposed, and one of these solutions would look as follows:

- add to  $sig(\mathcal{T}_{uni}^{**})$  both  $Employee_{vis}$  and  $Employee_{invis}$ ;
- modify  $\mathcal{T}_{uni}^{**}$  by adding to it the following axioms:

$$Employee_{vis} \sqsubseteq Employee \quad (173)$$

$$Employee_{invis} \sqsubseteq Employee \quad (174)$$

- rename in  $Map(\mathcal{T}_{uni}^{**}, \mathcal{T}_{com})$   $Employee$  to  $Employee_{vis}$ .

To interpret this repair, consider that instances of  $Employee$  in the repaired network may either be instances of  $Employee_{vis}$  or of  $Employee_{invis}$  or of  $PhDStudent$ . The first ones are visible to  $\mathcal{T}_{com}$  because they are mapped to  $Staff$  in  $\mathcal{T}_{com}$ . In other words, instances of  $Employee_{vis}$  in a university, like



staff members in a company, cannot be trainees. In the WMS-repaired network the classification of *bruce* does not yield a contradiction, nor does the classification of instances of other concepts. Although not strictly needed, an additional modification of WMS for DL may also repair the entailment between *PhDStudent* and *Employee*, and make the former a subconcept of *Employee<sub>invis</sub>*, i.e., university employees that, unlike staff members in a company, can be trainees.

## 9 Further work

In this paper we have described case studies of the application of three Ontology Repair Plans that evolve physical theories that are either inconsistent with the experimental evidence or are redundant. We have also evaluated our ORPs on some additional case studies that we do not describe in detail here. For instance, WMS has been tested on models of the hypothesis of the existence of a planet called Vulcan, which was meant to explain peculiarities of Mercury's orbit (today explained by general relativity). *Inconstancy* has been applied to models of the discovery of various laws of fluids, for instance, Boyle's law (in a closed system, the inversely proportional relationship between the pressure and volume of a gas depends on temperature).

There are many further potential case studies that we have not had the resources to implement. The New Scientist magazine, for instance, is a prolific source of such potential case studies, generating one or two candidates per week. The majority of these are case studies for the WMS ORP, with *Inconstancy* being the second most frequent ORP. The reason we cannot implement all these case studies is that formalising the area of Physics under-pinning them requires a significant investment of time and effort. If there were a pre-existing formalisation of all physical theories, then implementing such case studies would be much simpler and would facilitate deeper inference that would extend the range of ontology repairs. Unfortunately, only a few small fragments of the required formalisation currently exist. Developing such an all encompassing formalisation, to include not only all currently accepted theories but also those now disproved, would be an immense undertaking, requiring resources far beyond that available to our project.

We have designed and implemented additional ORPs that are not mentioned in this article. These include the following:

*Reidealisation*: Given two ontologies that disagree over the value of some measurement, which is the kind of conflict WMS detects and repairs, one could resolve a special case of this kind of conflict without inventing some invisible component and taking the viewpoint that the original conceptualisation gives only a partial view of the underlying property. Instead, the idealisation of the

property could be changed such that it is viewed as being a property of another type. With the new idealisation, the measurement function, which originally returns different values in each of the two ontologies, will no longer return conflicting values.

*Spectrum*: A variation of *Unite* is to realise that a number of apparently different objects are different manifestations of the same phenomena. The classic example of this ORP is the realisation that light, radio waves, X-rays, etc. are all kinds of electro-magnetic waves that lie along a spectrum.

We have also designed other ORPs but have not had the resources to implement or evaluate. These include the following:

*Constancy*: Just as *Unite* is a dual of WMS, *Constancy* is the dual of *Inconstancy*. Instead of revealing that a function has a hidden dependency, *Constancy* shows that a function is *independent* of an assumed dependency. Historic examples of *Constancy* include Galileo Galilei's demonstration that the acceleration of an object due to gravity is independent of its mass and his observation that the period of a pendulum is independent of its amplitude.

*Unify*: A more extreme version of *Spectrum* is to Unify two or more ontologies by replacing them with a new ontology that subsumes them all. The current search for a Grand Unified Theory to encompass both gravity and quantum mechanics is an attempt to apply *Unify*. Classic applications of *Unify* were Maxwell's unification of electricity and magnetism and Glashow, Weinberg and Salam's unification of electromagnetic and weak forces.

*Analogy*: An under-appreciated role of analogy is the initial population of an ontology for a new phenomenon by the recycling of an ontology for an already understood phenomenon. The classic example of the *Analogy* ORP is the initial population of an ontology for electricity by analogy with the ontology for hydraulics. Other popular applications of *Analogy* are the many attempts to adapt evolutionary theory to explain economic and social situations.

We have developed our theory of ontology evolution in the Physics domain largely because of the plentiful supply of well documented case studies on which to evaluate it. However, as our discussions in Figure 1 and Section 8.5 illustrate, we intend our theory to be a general one that is applicable to other domains and adaptable to other logics. In particular, we intend to continue our study of its application to Description Logics and its ability to extend the range of belief revisions to new kinds of plausible revisions. Furthermore, it would be interesting to conduct further experiments to distinguish between the plausible repairs and allow the heuristics to select the historically accurate solution.

## 10 Conclusion

In this article we have presented a Higher-Order Logical approach to ontology evolution in the domain of Physics. As argued in Section 2, the need for a higher order approach to the evolution of formal representations of concepts and their relationships (e.g., ontologies, knowledge bases, logical theories, etc.) arises from the inherent lack of generality of existing DL and FOL approaches. In order to go beyond axiom retraction and/or addition and achieve subtler forms of automated diagnosis and repair, a meta-logic more expressive than DL or FOL is required. We have shown that, especially because of the polymorphism of its symbols, HOL can be such meta-logic. HOL's formulae can abstract over types, number of arguments, etc. thus expanding the logical search space and providing elements for the formulation of *declarative theories* of ontology evolution – as opposed to the *algorithms* usually found in the DL- or FOL-based literature.

This higher-order approach under-pins a framework which includes both a methodology for the formalization of ontology evolution in HOL and an implementation of this methodology in the theorem prover Isabelle. The main elements of the GALILEO framework are:

*Ontologies formalised as (sub-)contexts by using locales*, i.e., as multiple logical theories.

*Bridges*, i.e., contexts that are used to resolve possible linguistic differences between ontologies that do not share exactly the same language.

*Conflicts between ontologies* that are each *locally consistent or parsimonious* but that *globally contradict each other or contain redundancies with respect to one another*. These conflicts are inspired by case studies in the history of Physics, in which clashes between theories, experiments, observations, etc. resulted in the field's evolution.

*Ontology Repair Plans*, i.e., mechanisms for the diagnosis of conflicts between locally consistent or parsimonious but globally inconsistent or redundant ontologies as well as for their evolution by means of axiom retraction and/or addition and/or signature change.

The presentation, application and evaluation of three of GALILEO's Ontology Repair Plans has provided empirical evidence of the value of a HOL-based approach to ontology evolution. In Sections 4 through 6 the ORP *Where's My Stuff?* was applied to the discovery of latent heat and to the postulation of dark matter; *Inconstancy* was applied to Modified Newtonian Dynamics in the study of galaxies and to the observations that proved the speed of light to be finite; *Unite* was applied to the identification of the Morning and Evening Stars and to the assessment of the shape of the Earth. This formal apparatus set the stage for the discussion and evaluation of the implementation in Sections 7 and 8.

The implementation is prototyped in the higher-order theorem prover Isabelle, using *Locales* to represent ontologies and higher-order deduction and matching with heuristics to perform diagnosis and repair. The prototype has proven to be a valuable tool for exploring the space of solutions. On the one hand, the power of HOL and HOM is unleashed on the case studies, and this results in a number of matches that is often huge. On the other hand, the heuristics have proven effective in curbing this power, by bringing the solution space to a manageable size, yet always returning the historical solution alongside other physically meaningful ones.

**Acknowledgements** The research reported was supported by grant EP/E005713/1 of the Engineering and Physical Sciences Research Council (EPSRC) and by an award of the Overseas Research Students Awards Scheme (ORSAS).

## References

- C. Ballarín. Locales and locale expressions in Isabelle/Isar. *Lecture Notes in Computer Science*, 3085:34–50, 2004.
- A. Borgida and L. Serafini. Distributed description logics: Assimilating information from peer sources. *J. Data Semantics*, 1:153–184, 2003.
- P. Bouquet and L. Serafini. On the difference between bridge rules and lifting axioms. In *CONTEXT*, pages 80–93, 2003.
- P. Bouquet, F. Giunchiglia, F. Van Harmelen, L. Serafini, and H. Stuckenschmidt. Contextualizing ontologies. *Web Semantics: Science, Services and Agents on the World Wide Web*, 1(4):325–343, 2004. ISSN 1570-8268.
- A. Bundy and M. Chan. Towards ontology evolution in physics. In Wilfrid Hodges and Ruy de Queiroz, editors, *Logic, Language, Information and Computation*, volume 5110 of *Lecture Notes in Computer Science*, pages 98–110. Springer Berlin / Heidelberg, July 2008.
- L. Cambresy, S. Derriere, P. Padovani, Preite-Martinez A., and A. Richard. An ontology of astronomical object types for the virtual observatory version 1.3. In *International Virtual Observatory Alliance Note*, 2010. URL <http://www.ivoa.net/Documents/Notes/AstrObjectOntology/>.
- M. Chan and A. Bundy. Inconstancy: An ontology repair plan for adding hidden variables. In S. Bringsjord and A. Shilliday, editors, *Symposium on Automated Scientific Discovery*, number FS-08-03 in Technical Report, pages 10–17. AAAI Press, November 2008. ISBN 978-1-57735-395-9.
- M. Chan, J. Lehmann, and A. Bundy. Higher-order representation and reasoning for automated ontology evolution. In *Proceedings of the 2010 International Conference on Knowledge Engineering and Ontology Development*,

- pages 84–93. SciTePress, 2010. ISBN 978-989-8425-29-4.
- A. Doan, J. Madhavan, P. Domingos, and A. Halevy. Ontology matching: A machine learning approach. *Handbook on Ontologies in Information Systems*, pages 397–416, 2004.
- J. Du, G. Qi, and Y.D. Shen. Lexicographical inference over inconsistent dl-based ontologies. *Web Reasoning and Rule Systems*, pages 58–73, 2008.
- G.F.R. Ellis and J.P. Uzan.  $c$  is the speed of light, isnt it? *American journal of physics*, 73:240, 2005.
- G. Flouris. On belief change in ontology evolution. *AI Commun.*, 19(4):395–397, 2006.
- G. Flouris, Z. Huang, J.Z. Pan, D. Plexousakis, and H. Wache. Inconsistencies, negations and changes in ontologies. In *Proceedings of AAAI*, pages 1295–1300, 2006.
- S. Ghilardi, C. Lutz, and F. Wolter. Did I damage my ontology? A case for conservative extensions in description logic. *Proceedings of the Tenth International Conference on Principles of Knowledge Representation and Reasoning*, pages 187–197, 2006.
- F. Giunchiglia and P. Shvaiko. Semantic matching. *The Knowledge Engineering Review*, 18(03):265–280, 2004.
- M. J. Gordon and A. M. Pitts. The hol logic and system. In J. Bowen, editor, *Towards Verified Systems*, chapter 3, pages 49–70. Elsevier Science B. V., 1994.
- P. Haase, F. van Harmelen, Z. Huang, H. Stuckenschmidt, and Y. Sure. A Framework for Handling Inconsistency in Changing Ontologies. In *The Semantic Web-ISWC 2005: 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland, November 6-10, 2005: Proceedings*. Springer Verlag, 2005.
- M. Homola. Distributed description logics revisited. In *Description Logics*, 2007.
- F.V. Jensen. *An introduction to Bayesian networks*, volume 36. UCL press London, 1996.
- Q. Ji, P. Haase, G. Qi, P. Hitzler, and S. Stadtmüller. RaDON – Repair and Diagnosis in Ontology Networks. *The Semantic Web: Research and Applications*, pages 863–867, 2009.
- A. Kalyanpur, B. Parsia, E. Sirin, and B.C. Grau. Repairing unsatisfiable concepts in OWL ontologies. In *ESWC*, pages 170–184. Springer, 2006a.
- A. Kalyanpur, B. Parsia, E. Sirin, B.C. Grau, and J. Hendler. Swoop: A web ontology editing browser. *Web Semantics: Science, Services and Agents on the World Wide Web*, 4(2):144–153, 2006b.
- R. A. Kowalski and M. Sergot. A logic-based calculus of events. *New Generation Computing*, 4:319–40, 1986.
- J.S.C. Lam, D. Sleeman, J.Z. Pan, and W. Vasconcelos. A fine-grained approach to resolving unsatisfiable ontologies. *Lecture Notes in Computer Science*, 4900:62, 2008.
- P. Langley, J. Zytkow, H. Simon, and G. Bradshaw. Mechanisms for qualitative and quantitative discovery. In R. S. Michalski, editor, *Proceedings of the International Machine Learning Workshop*, pages 121–132. University of Illinois, June 1983.
- U. Lösch, S. Rudolph, D. Vrandečić, and R. Studer. Tempus fugit. In *Proceedings of the 6th European Semantic Web Conference on The Semantic Web: Research and Applications, ESWC 2009 Heraklion*, pages 278–292, Berlin, Heidelberg, 2009. Springer-Verlag. ISBN 978-3-642-02120-6. doi: 10.1007/978-3-642-02121-3\_23. URL [http://dx.doi.org/10.1007/978-3-642-02121-3\\_23](http://dx.doi.org/10.1007/978-3-642-02121-3_23).
- F. McNeill and A. Bundy. Dynamic, automatic, first-order ontology repair by diagnosis of failed plan execution. *International Journal On Semantic Web and Information Systems*, 3(3):1–35, 2007. Special issue on ontology matching.
- T. Nipkow, L. C. Paulson, and M. Wenzel. *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*, volume 2283 of LNCS. Springer, 2002.
- L. C. Paulson. *Isabelle: A generic theorem prover*. Springer-Verlag, 1994.
- V. C. Rubin, N. Thonnard, and Jr. Ford, W. K. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 ( $R = 4\text{kpc}$ ) to UGC 2885 ( $R = 122\text{ kpc}$ ). *Astrophysical Journal*, 238:471, 1980. doi: 10.1086/158003.
- S. Schlobach and R. Cornet. Non-standard reasoning services for the debugging of description logic terminologies. In *International Joint Conference on Artificial Intelligence*, volume 18, pages 355–362. Citeseer, 2003.
- H. Stuckenschmidt. Implementing modular ontologies with distributed description logics. In *WoMO*, 2006.
- H. Stuckenschmidt, F. van Harmelen, P. Bouquet, F. Giunchiglia, and L. Serafini. Using c-owl for the alignment and merging of medical ontologies. In Udo Hahn, editor, *KR-MED*, volume 102 of *CEUR Workshop Proceedings*, pages 88–101. CEUR-WS.org, 2004. URL <http://dblp.uni-trier.de/db/conf/krmed/krmed2004.html#StuckenschmidtHBGS04>.
- M. Wisner and S. Carey. When heat and temperature were one. In A. Stevens and D. Gentner, editors, *Mental Models*, pages 267–297. Erlbaum, 1983.